

Does biotechnology have a price?

The five United States universities represented at the meeting to consider the commercialization of genetic manipulation have issued a statement whose omissions are more significant than its content.

Last week's gathering of five university presidents and their entourages at Pajaro Dunes in California was a predictably civilized occasion it would seem, but hardly productive. The communique afterwards put out (see *Nature* 1 April, p.381) delicately acknowledges that the interest of universities in the commercial exploitation of genetic manipulation may occasion difficulties; makes a few general observations to the effect that universities should not let corporate or individual relationships with external commercial interests interfere with teaching and research; says that universities holding patents should be free occasionally to award exclusive licences; and concludes that it will in most circumstances be unwise for universities to invest financially in companies in which faculty members have a direct and substantial interest.

These unremarkable observations are offered as a contribution to what the five presidents rightly expect will be a continuing discussion of the propriety of arrangements between universities and external commercial interests. Because groups of university presidents are free to spend a few days at the seaside in each other's company without issuing an account of what they talked about, the Pajaro Dunes statement is something to be grateful for. Indeed, the fact of the meeting implicitly and the communique explicitly acknowledge that the public interest in the commercialization of academic research overrides the interests of individual universities. So why does the communique fail even to mention the immediate and perplexing issues with which universities must grapple in the years ahead?

Take secrecy. One of the most obvious consequences of the interpenetrating networks of commercial agreements being forged with external commercial interests by institutions, departments and individual academics is that the free flow of ideas within laboratories will be inhibited by the fear that to say too much will be to hazard a commercial advantage. The Pajaro Dunes communique acknowledges the risk, and says that agreements with external interests should be "so constructed as not to promote a secrecy that will harm the progress of science. . ." But how, in the real world, is that to be accomplished?

What happens, for example, if two members of the same academic department are also the prime movers in two separate commercial companies? Or if two close colleagues are consultants for competing commercial concerns? Or if the research carried out by one group within a laboratory is supported by a commercial organization by means of an agreement that gives the company first refusal of patent rights? In each case (and others), close colleagues will be tempted to bite off their tongues for fear of providing clues revealing of the commercial interests of their sponsors. Previously innocent academics, probably less skilled than professional consultants at distinguishing between proprietary and general information, are likely to behave like clams even in discussions of such general questions as the mechanisms by which eukaryotic genes are regulated. For the time being, in any case, (the kinds of questions that commercial companies are likely to be asking of those who work for them as consultants.) But in most universities and academic departments, the tongues of externally-interested academics will be further tied by the knowledge that they cannot know which of their colleagues is working privately for which competing organization.

Disclosure of external interests is therefore an essential part of

the process of preserving the civility of those parts of the academic research enterprise recently discovered to be commercially important. But disclosure to whom? And to what extent? Will it suffice that the chairman of a department should know for whom his immediate colleagues are moonlighting? Will chairmen in their turn shoulder the responsibility implied in the confidentiality of the disclosures made to them of making sure that putatively competing academics are aware of each others' outside interests? Or might it be better that disclosure should be more general? And is there not then a danger that students would be apprised of information that would be an excuse for occupying the departmental chairman's office on some sunny afternoon when there is nothing else to do? It is no wonder that the participants in the meeting at Pajaro Dunes steered clear of questions such as these in their published statement. Questions such as these are nevertheless prominent among those universities that must decide for themselves in the next few weeks or months.

The Pajaro Dunes communique would have been more helpful if it had openly acknowledged that problems such as these exist. It should also have recognized that the difficulties occasioned by biotechnology differ only in degree from those with which universities have learned, however uneasily, to live. For biotechnology, the commercial exploitation of genetic manipulation differs for the time being from more familiar branches of technology only in the speed with which a bright idea may be translated into a patentable invention. In other fields, chemical engineering for example, the process of innovation consumes more time and is less certain. But it is well understood that much of the research carried out in university laboratories, in computer science for example, has led in recent years to important industrial improvements and, as a consequence, to a close interdependence between academics and industrial partners. The result, especially in the United States, is that there is now a modest army of academics acting as consultants to industrial companies while agreements between companies and academic laboratories for the support of basic science in return for beneficial treatment in the exploitation of innovations are by no means unprecedented or unwelcome.

From this it follows that universities cannot hope to devise rules for the proper exploitation of basic research in genetic manipulation unless these also apply to the regulation of other branches of technology. It is, however, plain that in biotechnology, more explicit rules for deciding how academics and academic departments should conduct themselves are necessary. Within departments, seemingly rules about disclosure are plainly needed if academics are to continue to talk to each other without restraint. For universities, it is also essential that the basis on which agreements for the support of research by outside companies should be publicly disclosed. Failure in this respect will put in hazard public trust in the whole of higher education. And universities must more zealously than at present ensure that the education of students, graduate students in particular, is not distorted by the commercial extramural interests of those who are supposed to teach them. If this implies that the research projects on which graduate students are engaged should not be determined exclusively by individual members of a university faculty, the restriction of individuals' academic freedom that would imply would be a small price to pay for the preservation of the



reputation of universities to put the interests of their students first. Individual universities may not be too hard-pressed to devise rules for regulating what happens in departments with an interest in biotechnology. The difficulty will be to ensure that these can be applied across the board.

Some thought should also be given to the question of why the commercialization of basic research has recently thrown up so many problems. Government support for basic research is shrinking everywhere, while universities are being urged to look to industry for sponsorship. Universities and individual academics are being driven into the arms of those companies farsighted enough to recognize that they have much to gain from partnerships like these. There is much to be said for finding ways of turning academic discoveries into prosperity for the wider community. But while some academics appear genuinely to be excited by the challenge of turning their bright ideas into business enterprises, others appear to be at least as much intrigued by the prospect of monetary rewards. Explanations, or excuses, are not hard to find. Compared with their students who make a success in industry, academics are scandalously paid, for example. But has the traditional loyalty of academics to the institutions which employ them been eroded? And if so, what will be the consequences of that? It is a fair guess that the university presidents who met at Pajaro Dunes two weeks ago would have been fully conscious of this disturbing thought.

The Falkland War

The British Government hopes that diplomacy will avoid a war with Argentina. How?

The next few days are bound to have an air of unreality not merely for those who live and work in Britain or Argentina but for the rest of the world. A small British naval force is sailing to the South Atlantic with the intention of liberating the Falkland Islands from their occupation by an Argentinian force last weekend. The military task is formidable, for it is bound to be more difficult to dislodge an occupying force than it would have been to prevent it from landing in the first place. So too is the diplomatic task, for the government of Argentina has repeatedly declared its intention of staying put. Sensibly, however, the British government has promised to use the next few days for seeking some kind of solution. What might this be?

The facts are clear. There is no impartial body of opinion that disputes the illegality of what the Argentinian government has done. Moreover, the doctrine advanced in Buenos Aires that the Falkland Islands belong to Argentina because of their proximity is a dangerous doctrine, as the case of Ulster in Ireland shows. The Argentinian action in the Falkland Islands might have been seen in a different light if the 1,800 inhabitants of the islands had shown the slightest inclination to opt for Argentinian rather than British citizenship, and if Britain made no use of the territory (which serves as a staging post not merely for commercial vessels but for British Antarctic expeditions). In this sense, the British presence in the Falkland Islands is beyond reproach. The problem is to reconcile this with the apparently implacable conviction of the government of Argentina that the Falkland Islands belong to Argentina.

Sovereignty is a heady concept, for which people are all too willing to kill each other. In the weeks ahead, the governments of Britain and of Argentina will also know that their own survival may depend on how things turn out in the Falkland Islands. So should they not think of trying to reach a settlement of this wasteful conflict by more old-fashioned methods than the use of military technology. Specifically, should not Britain think of offering to sell the Falkland Islands to Argentina in much the way in which Alaska was traded by the Soviet Union to the United States? The price, of course, would be high. Not merely would the 1,800 people of the Falkland Islands have to be compensated, but the British government would have to be recompensed for the loss of the benefits it now enjoys from possession of this out-of-the-way place as well as for the benefits that may in the future accrue.

Rewiring Britain

The Government, after a decade of indifference, is now too quickly embracing cable television.

As Paul on the road to Damascus, the British government has suddenly been converted to the belief that it will be socially beneficial if a substantial part of Britain can be covered with cable television systems. Two weeks ago (see *Nature* 25 March, p.282) the Department of Industry surprised most British taxpayers with an announcement that by the end of the year, it will have worked out a scheme that will allow commercial companies to invest in cable television systems. Between now and then, the government hopes, a committee under Sir John Hunt, one of the many ex-secretaries of the Cabinet still in active service, will have worked out a set of rules for regulating this business, new for Britain. No doubt the government has it in mind that many voters at the next general election will go more cheerfully to the polls if they are able to return to a wider choice of television signals than at present.

The directness of the government's change of heart is remarkable. For the past decade, the Home Office has allowed only a few local experiments with cable television. Part of its reluctance is explained by its wish not to encourage competition with the development of the national broadcast television service. Thus, in the autumn and after years of argument, a fourth channel of broadcast television will take the air financed by the revenues from commercial advertising. Hitherto, the government has listened to the pleas of the Independent Broadcasting Authority that the success of that enterprise should not be jeopardized by allowing advertisers competing outlets for their cash. That principle, it seems, has now been abandoned. At the same time, the present British government has plainly decided that it can no longer allow itself to be restrained by the squeamishness of the principal opposition party about pay-television. Subscribers to the proposed cable television networks will be able to watch on their screens whatever they can afford to pay for.

But what? This is where Sir John Hunt's committee will run into trouble. Working against the government's clock, the committee will be tempted to follow precedents established overseas, in the United States especially. It will thus be tempted to insist that any new cable system should distribute the four national television channels that are at present adequately served by conventional broadcasting. Even though some of the would-be operators of cable systems are likely to protest that such a requirement would be pointless, in the long run substantial benefits should accrue. The other obvious temptation will be to follow the common pattern in the United States which requires cable systems to provide one or more channels free to those local interests believing they have something to say to the public at large. That is a temptation to be resisted, for the experience of the United States has shown that people linked with cable systems vote with their tuning knobs to watch something else in preference. In the long run, the public interest will be best served by taking its chance with other interests.

The most contentious issue, however, will be that of how the content of signals put out on the new cable systems should be regulated, and by whom. There will be a temptation to set up some public body to ensure, for example, that nothing broadcast by the new cable systems can give offence to those who at present act as custodians of public morality. This, too, is a temptation that should be resisted. For the most obvious difference between cable systems and conventional broadcast channels is that the channel capacity of cable is much greater than that of the atmosphere. While it will remain important that the national television service should conform with rules that are understood and accepted, people linked with cable television should be allowed to receive whatever signals they wish — and which they are prepared to pay for. The danger here is that the potential operators of cable networks, anxious as they are to encourage the government in its new resolve, will meekly agree to censorship that will not afterwards be easily removed.

German cancer centre criticized

Report urges major changes in constitution

Sweeping changes at the Deutsches Krebsforschungszentrum (DKFZ) in Heidelberg have been recommended by the international commission appointed by the German Minister for Science and Technology. The proposed changes would invest the director of DKFZ with executive power, the lack of which caused the resignation of Professor Hans Neurath last year (see *Nature* 293,252; 1981).

The report, published last week by the ministry, is highly critical of the overall performance of DKFZ in basic research. Although acclaiming the quality of research by some investigators in the centre's eight institutes, particularly those of cell biology, immunology and virology, the report says that in a large number of the institute's 39 divisions, the research work is merely reliable and unimaginative. And the output of a number of scientists has amounted to little more than one unremarkable paper a year for the past decade: Overall "this does not indicate that the quality of research commissioned in DKFZ is in any way outstanding".

The centre is also criticized for its lack of collaborative ventures and its failure to take a lead in arranging multi-centre clinical trials. The commission considers that the largest cancer centre in Western Europe, employing more than 1,000 staff at an annual cost of DM90 million (£20 million), should do better.

Two main factors are said to account for the mediocre performance of DKFZ. The first is lack of effective external peer review. The second is that the constitution of the centre is such that the director does not necessarily have the power to bring about changes. "His proposals may be blocked or intolerably delayed and his position can be undermined by other proposals . . . from the Scientific Council" (a committee made up of the heads of institutes and an equal number of elected staff members). It was intolerable difficulties of this kind that Professor Neurath had "bravely attempted to overcome", says the report.

The resolution of this problem, says the commission, is of paramount importance to the future of DKFZ. It proposes that DKFZ's director (officially chairman of the executive board) should also serve as chairman of the Science Council. In this way the director could no longer be subordinate to his staff and could assume authoritative direction of the centre.

In the commission's view, the Science Council would still be a powerful advisory

group whose majority view its chairman, the director, could ignore only at his peril (his is the only senior post without tenure).

On peer review, the commission says there is no adequate system to ensure that resources are wisely distributed among a staff most of whom have unlimited tenure. Some form of peer review has periodically been carried out by the centre's Science Advisory Council (Wissenschaftliche Beirat) but the council is not required by the constitution to conduct such reviews, is too small to perform the job adequately and is open to influence from the staff.

The commission proposes that the Science Advisory Council be replaced by a committee of eight independent scientists who would be responsible for arranging regular external peer reviews of all institutes and divisions of DKFZ. The same independent scientists would also fill the eight places for scientists on the Kuratorium. No longer would three of these posts be filled by elected members of

DKFZ. The commission considers that staff members have no place on the executive board of the institute that employs them whereas the director, who is not a voting member of the Kuratorium, should be on the board.

It is on the question of "worker participation" that there is bound to be political opposition to change. Staff representation on the Kuratorium began in 1975, when it was being widely adopted in both public and private concerns throughout Germany. Dr Wolfgang Finke, now chairman of the Kuratorium, would prefer to see it retained on the grounds that the presence of staff scientists on the Kuratorium has been a useful channel of information between the staff and the trustees. If there is a problem, he feels, it arises from the individuals concerned, not the principle.

Apart from that point, Dr Finke believes that the Kuratorium should give serious consideration to all of the commission's

Yale says no to a \$33,000 grant

Washington

Yale University last week reluctantly bit one of the hands that feed it, and declined a \$33,000 grant offered by the National Science Foundation for one year beginning on 1 May. By this unprecedented act, the university also gave renewed publicity to the fact that one of its faculty, mathematics professor Dr Serge Lang, strongly objects to the US government's requirement that recipients of federal research grants should produce detailed reports of the amount of time they spend working on federally supported projects.

Dr Lang described himself on the telephone last week as one "who knows more about A21 than anybody else". The reference is to the much disputed circular A21, issued by the Bureau of the Budget (now the Office of Management and Budget, OMB) more than twenty years ago. Yale's refusal of the grant, the renewal for a third year of its support for Dr Lang's project "Number theory and elliptic curves", is due to his refusal to accept the reporting requirements.

Dr Lang also said last week that his objections to "effort reporting" were unlikely to be met by the revision of A21 now being considered by OMB, which are substantially those agreed last year between OMB and groups such as the American Association of Universities and the Council of Scientific Society Presidents.

Since the publication of the proposed revision in the *Federal Register* on 7 January, there has been a steady flood of protests from academic scientists that even the proposed revisions do not meet the objection that it is meaningless to attempt

to distinguish between time spent on teaching and time spent on research. As published, the revision of A21 would do away with the requirement for a detailed account of the time spent on various activities by faculty members supported partly by federal funds. Instead, three alternative methods of accounting for effort devoted to federal projects are proposed, one of which would allow of a certifying signature by "a person having direct knowledge of the work" and not by the faculty member as such.

In the draft revision, however, OMB has not conceded the principle urged by many academic scientists that there are circumstances, research by graduate students for example, in which it is literally meaningless to seek to distinguish between research and teaching. Opponents of OMB's accounting ambitions are able with relish to quote on this point from the report of a task force in 1968 under Cecil E. Goode, then an official of the Bureau of the Budget.

Part of the reason why A21 has become a source of contention in the past three years is that government auditors have been seeking to use it as a means of quantifying (and presumably limiting) the indirect costs added to research grants as compensation for the administrative work carried out centrally by a university. Yale was not impelled to decline the first two instalments of Dr Lang's three-year grant because of a dispensation which expired at the beginning of this year.

Under the terms of the grant offered, Yale would have been paid two-ninths of Dr Lang's salary by way of compensation for work done during the long vacation. ●

recommendations. He will no longer be its chairman when the Kuratorium comes to consider the future of DKFZ on 21 June because, from 1 May responsibility for the centre will be shifted as part of a reshuffle of responsibilities in the ministry. It is likely that the new chairman will be Dr Güntsch.

Professor Hans Neurath, now back in the University of Washington, Seattle, also welcomes the report, describing it as "excellent, clear and succinct".

Reaction to the report at DKFZ is hard to gauge, but inevitably mixed.

Only after the 21 June meeting of the Kuratorium will it be known which of the commission's proposals will be adopted. One problem is that since Professor Neurath's resignation, DKFZ has been without a real director, although Professor Otto Westphal was appointed acting director from 1 March until the end of 1982. Another is that there is almost certain to be strong resistance to many of the proposals from within the centre. And it does not help that the politicization of the centre's difficulties comes at a time when the German government has matters of far greater importance on its plate — survival for example.

Thus, whereas the commission recommends that proper peer review of DKFZ starts without delay, with at least three institutes to be reviewed by the end of the year, there is a good chance that the peers will need topcoats rather than safari suits by the time they descend upon Heidelberg.

Peter Newmark

Europe's nuclear power

Border incidents

Brussels

A bizarre mixture of Molotov cocktails and red and yellow balloons was released at a demonstration on 27 March when Belgian environmentalists protested at the continuing construction of four French nuclear power stations at Chooz, close to the Belgian border. Protests against the Chooz reactors are now a regular part of Belgian life. Some weeks ago, the Belgian Embassy in Paris was occupied by two Belgian senators belonging to the environmental party, and demonstrations on the border are now promised on the last Saturday of each month.

While there is no sign that the protests will influence the French government's determination to press ahead with the four 1,300 MW reactors to be built on the Meuse at Chooz, they may affect the Belgian government's willingness to participate in the project. Given the prospect that Belgium will have to rely on nuclear power for half of its electricity production by 1985, the government must either sanction the nuclear power construction in Belgian or join in the Chooz project.

For a time, the government sought to obtain electricity and to counter the protests by linking participation at Chooz

with the stipulation that the French reactors should be built to Belgian safety standards. This suggestion has, however, been rejected by the French on the grounds that construction costs would be increased.

The issue has thus been invested with national pride. The Belgian government is being told, most vociferously by the Flemish-speaking population near the border, that it has allowed itself to be bullied by its larger neighbour. France, on the other hand, cannot admit that its own safety standards for reactors are inadequate simply because the Belgian standards are more stringent. France has also declined to reveal details of the reactors to be built at Chooz beginning in December, and to reveal plans for the discharge of some low-level waste into the Meuse.

There is a chance that the European Commission may be able to help. The revision of Article 37 of the Euratom Treaty soon to be adopted would require France to seek approval for its plans for radioactive waste disposal six months before construction at Chooz begins. So far, however, there is no sign that France is prepared to let the commission interfere or to engage in consultations with Belgium.

The Belgian government is in any case itself divided over the country's energy plans. The energy minister, Mr Etienne Knoop, is anxious that Belgium should have more nuclear power stations and keen on participation at Chooz. He is opposed by the science and budget minister, Mr Philip Maystadt, who wants the decision on Chooz put off until the parliamentary debate arranged for the end of May — after the 16 April deadline for a decision on participation will have passed.

In the end, Maystadt's arguments are likely to prevail. Belgian utility companies argue that it would be more sensible to invest in a Belgian project than in Chooz, especially because it is now estimated that Belgium will need to build thirteen 1,300 MW reactors by the year 2020. In any case, the cost of electricity from the Chooz plant quoted by the French is unpalatably high.

Jasper Becker

●Environmentalists are also on the move in France — on foot, in fact, between the site of the near-complete French fast breeder reactor, Superphénix, in the south of France, to Paris, where they hope to meet President Mitterrand on 18 April to protest against fast breeder development. A hundred demonstrators gathered on Sunday near Superphénix in the march, which has begun peacefully. This is in marked contrast to the rocket attack on Superphénix last summer, and the violent demonstrations of 1977 in which one demonstrator died. Brice Lalond, national secretary of the French Friends of the Earth, addressed the marchers. "Nothing has changed since 10 May", he said, referring to the election of Mitterrand as President of France and his socialist government.

Robert Walgate

United States astronomy

High priorities

Apparently unabashed by the austerity now rife in Washington, the long-awaited Field report on astronomy and astrophysics in the 1980s will include a \$1,900 million (at 1980 prices) shopping-list, mostly for new equipment. A draft of the report (to be published next month) says that these plans would require that the astronomy budget of the National Science Foundation (NSF) in the 1980s should be 30 per cent greater in real terms than in the 1970s.

The report has been prepared by the Astronomy Survey Committee under the chairmanship of the director of the Harvard-Smithsonian Center for Astrophysics, George Field. The committee was set up in 1978 by the National Academy of Sciences (NAS) to assign priorities to a wide range of astronomical projects based on the ground or in Earth orbit, but excluding planetary missions and deep-space probes.

Field report recommendations

	Decade cost (1980 \$ million)
Major new programmes, in order of priority	
1. Advanced X-ray Astronomy Facility (AXAF)	500
2. Very Long Base-line (VLB) Array	50
3. New Technology Telescope (NTT)	100
4. Large Deployable Reflector in Space	300
	950
Modest new programmes, in rough order of priority	
1. Augmentation of Explorer satellite programme	200
2. Far UV spectrograph in space	150
3. Space VLB antenna	60
4. Optical/IR telescopes (2.5 m)	20
5. Advanced Solar Observatory in space	200
6. Cosmic ray experiments	100
7. Search for Extraterrestrial Intelligence	20
	750
Small new programmes	
10 m submillimetre-wave radio antenna	4
Other important programmes	
Spatial interferometer for mid-IR	3
High precision optical astrometry programme	3
Temporary programme to maintain astronomical expertise at universities	10
	20
Total expenditure over decade on new project	1,720
Total expenditure on "prerequisites" (theory, data analysis, etc.)	190
Grand total	\$1,910 million

In its review of astronomical problems and the best ways of tackling them, the report assumes the existence of such shuttle-dependent projects as the Space Telescope, the Gamma-Ray Observatory, the Shuttle Infrared Telescope Facility, the Solar Optical Telescope and the ground-based 25-metre millimetre-wave radio-telescope, all but the last of which appear to have survived in the US Administration's budget for 1983.

For the 1980s, the committee recommends that the highest priorities should be given to the Advanced X-ray Astrophysics Facility (AXAF) satellite, a Very Long Baseline (VLB) array of radiotelescopes, a

giant New Technology Telescope (NTT) in the 15-metre class and a 10-metre sub-millimetre-wave antenna. All these devices would contribute to the solution of the astronomical problems highlighted in the report — the large-scale structure of the Universe, the evolution of galaxies, violent events (such as supernovae and processes in active galaxies and quasars), the formation of stars and planets and solar and stellar activity.

In the committee's view, AXAF is a must at a cost of \$500 million. The VLB has beaten the 15-metre telescope into second place because it would require considerably less preliminary development and would have by far the best angular resolution of all techniques — 0.3 milliarc seconds — possibly allowing the resolution of the nuclei of active galaxies and quasars at radio wavelengths. But because of the tremendous potential light-gathering power of the 15-metre telescope, an order of magnitude larger than present telescopes, the committee puts that on as high a level of scientific desirability as any other project examined.

The committee also urges increased support for what it calls prerequisites — instrument design (including a request for declassification of certain infrared detectors), theory, data analysis, computing facilities and laboratory astrophysics. To handle the data from AXAF, the committee recommends that a special institute be set up in support of the satellite, analogous with the Space Telescope Institute. As a consequence of its recommendations, the committee sees the need for an increase in personnel at all levels.

In his foreword to the report, Herbert Friedman, chairman of the Mathematical and Physical Sciences Assembly of NAS, stresses the need for an increase in the relative contribution of ground-based astronomy and, consequently, the NSF contribution. He also says that the technology of new developments in astronomy is dependent on but will also contribute to areas of science and to industrial and military applications.

In the present climate, it is hard to see the astronomers having many major desires fulfilled. Of the major recommendations of the influential Greenstein report, the 1972 predecessor of the Field report, only one — the VLA radiotelescope — was implemented in the less austere 1970s, for example. The Field committee emphasizes, however, that the level of expenditure it requests is equivalent in real terms to that which followed the Greenstein report. Moreover, Dr Friedman points out that they are hoping "that the purposes of the present fiscal policies will be achieved in a reasonably short period and that a healthier base of federal scientific support will then be restored". But if the Greenstein report can be viewed as a precedent, X-ray astronomers have struck lucky.

Philip Campbell

High-energy physics

Tunnel vision

A legal obstacle to the construction of LEP, a European particle accelerator which should put European physics in the 1990s well ahead of the United States, has been removed. But the shouting is not yet over.

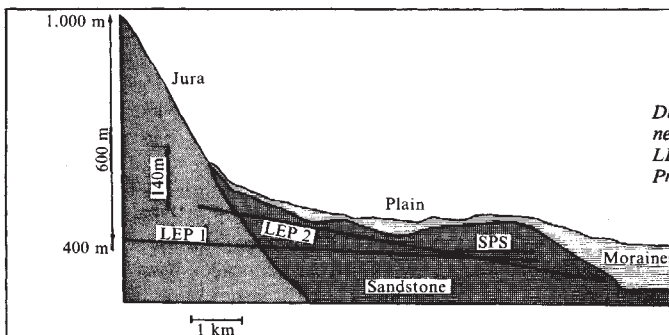
LEP — a large electron positron collider — will straddle the French-Swiss border in the plain of Geneva, linking with existing machines at the European Centre for Nuclear Physics, CERN. CERN has won cash and approval for the project from its member states, which include France and Switzerland, but the problem lately has been how to carry the project through local planning procedures in France and Switzerland. French environmentalists had managed by a legal technicality to block construction of an important exploration tunnel, but now the French court of appeal, the *Conseil d'Etat*, has finally ruled that the tunnel can go ahead after all.

This, however, is only a minor victory. The main justification for the tunnel was that it would provide an escape route for both engineers and water if the main LEP ring — which originally led right under the 1,000 m high peak of the Jura mountains — encountered high pressure water in the Jura limestone. But now the LEP ring has been shuffled almost entirely out of the Jura, by tipping it on its side (see figure), so water pressures will be lower. It will also now be possible to intervene from the surface if people are trapped.

However, there are still geological and political problems to solve. Geologically, the 3 km of LEP that will still lie in the Jura limestone lie entirely in the *piémont*, near the boundary with the sandstone of the plain. This rock is likely to be traversed by caverns and faults, and although vertical borings along the new route have revealed nothing serious, uncertainty persists.

Politically, the obstacle may still be the *Association gessienne de protection de la Nature*, the local French environmental organization presided over by M. Jean-Roger Honorat, the mayor of Echenevex. This group blocked the construction of the exploration tunnel, and although CERN has wooed it ever since, it remains suspicious.

The group will play an active part in the French planning process now under way.



New cancer research award

The 1982 Bristol-Myers Award for Distinguished Achievement in Cancer Research has been awarded to Dr Denis Burkitt (of Burkitt's lymphoma) and Dr Michael Epstein (of the Epstein-Barr virus). Their \$50,000 award was presented to them in New York this week.

Dr Burkitt was Government Surgeon in Uganda some twenty years ago when he proposed, from epidemiological studies, that an infectious agent caused childhood lymphoma in Africa. Dr Epstein heard Burkitt speak in London



Epstein (left) and Burkitt

in 1961 and within two years had isolated the infectious virus that causes the lymphoma.

Burkitt is now an honorary senior research fellow at St Thomas's Hospital in London. Epstein is professor of pathology at the University of Bristol. And Yvonne Barr, in case you were wondering, is an Australian housewife.

Peter Newmark

The French ministry of foreign affairs, with which CERN must deal, now has CERN's final environmental impact statement, and under French planning law will act on CERN's behalf. There is to be a local public inquiry at which anyone in the region may present his views. CERN itself will be represented at the inquiry only if needed to explain technicalities; it has no right to plead. Eventually the results of the inquiry and files of submissions will be returned to Paris for a decision by the *Conseil d'Etat*.

Meanwhile, CERN is about to invite tenders for the construction of the tunnel, which should be in by the autumn. A decision from the *Conseil d'Etat* is not expected before the end of this year but a delay beyond that would mean LEP would not be producing particle collisions by the end of 1987 as planned.

Robert Walgate

Agricultural Research Council

Savings plans

The British Agricultural Research Council (ARC) is still in a fix about its budget for the next three years. Plans to find all the necessary economies to meet the expected shortfall in its government grant at the Animal Breeding Research Organization (ABRO) and the Long Ashton Research Station have been abandoned. A decision on the future of the two institutes, which was to have been taken at the end of last month, has been put off to the end of April, and the council has set up a "core group" to look for comparable savings elsewhere.

ARC had originally planned to save the expected £3 million deficit in its annual budget by cutting research on food, beverages and fruit at Long Ashton and by cutting back ABRO to fundamental animal genetics. It has, however, found its hands tied by the Rothschild customer-contractor principle, whereby the Ministry of Agriculture (the customer) commissions a substantial proportion of research in ARC institutes (the contractors). The ministry, aided by the livestock industry, seems to have objected strongly to the plan to cut ABRO to one-fifth of its present strength. Such is the colour of the ministry's money, however, that at one stage it had almost persuaded ARC to cut fundamental rather than applied research. In the event, ARC last week produced a revised plan that may get it off that hook.

The revised plan being circulated for consultation before a final decision on 20 April entails a 50 per cent cut in ABRO's annual budget, from £2.25 million to £1.2 million, by 1984-85, but keeps the balance between fundamental and applied research roughly the same as now. Work on the long-term breeding of sheep and pigs and on the efficiency of dairy cows would continue, as would several areas of animal genetics. ARC says that it may also find a further £0.4 million for research in molecular biology later on. The council clearly hopes its new plan will have an easy passage.

The proposals for the Long Ashton Research Station, which have caused less of a stir, are likely to stand. Hence some research will probably be moved to the East Malling Research Station and the fruit industry is being asked to increase its support for the remaining programmes.

The debate over the future of the two institutes nevertheless has wider implications. ARC's own staff has complained that the process of regularly reviewing research programmes is not sufficiently open. It is also clear that the council needs some way of deciding priorities within as well as between institutes. But the most telling lesson of the past few days is the council's discovery that under the customer-contractor principle, the customer is ultimately (if not always) right.

Judy Redfearn

Not enough toxicologists in Europe?

The European Science Foundation (ESF) and the European Medical Research Council (EMRC) have announced a new programme for research and training in toxicology. A number of awards will be made in 1983 along the lines of ESF's existing brain research programme.

EMRC is basically a talking shop for ESF, representing members of the governing bodies of 14 Western European medical research councils who meet twice a year, and directors of the regional office of the World Health Organization and of the National Institute of Health in Bethesda, Maryland have a standing invitation to attend meetings. Toxicology has been a major interest of EMRC since it was established in 1971, and now nine members of EMRC have put up the money for the toxicology programme; the 1982 budget is FF700,000 (£63,600).

Although the demand for toxicologists in industry and government is growing rapidly, toxicology remains an ill-defined subject which may mean different things in France and, say, Denmark.

EMRC has recommended that priority should be given to mechanisms of toxicity and the epidemiology of exposure to chemicals. It also emphasizes the need for more research in toxicology and the inclusion of toxicology in the university curriculum. In the United Kingdom, for example, toxicologists are drawn mainly from pathology and biochemistry, and the subcommittee report on which the ESF programme is based is impressed with the complaint from industry that university graduates lack practical training in toxicology.

As well as travel and subsistence grants for training courses, ESF plans to award one-year training fellowships to graduates from toxicology-related disciplines. There will also be renewable fellowships for people already working as toxicologists for projects involving collaboration with teams abroad.

ESF is optimistic about the programme. It is to be hoped that the length of time it has taken to set up the new programme is not indicative of the pace at which all collaborative European research will develop.

Jane Wynn

Soviet science and technology

Practical slant

The Soviet Academy of Sciences seems to be planning a more direct interest in production research, according to Dr Anatolii Aleksandrov, the president, in his address to the academy's annual meeting. The academy's increased participation in technological projects in the past eighteen months seems to be connected with the appointment of Dr Gurii Marchuk, formerly chairman of the Siberian branch of the academy, as head of the State Committee for Science and Technology.

When Dr Marchuk moved to the state committee, which is responsible for applied research and its implementation, he urged greater use of the academy's talents in implementing research results in production — thus bridging what has become a major planning gap.

The academy's new practical slant is clearly approved by Dr Aleksandrov, who has often stressed that the Soviet Union is potentially independent of foreign research and development. By contrast, the chief academic secretary of the academy, Dr Georgii Skryabin, told the annual meeting of his regret that the deterioration of the international situation has led to reduced contact with the West. He particularly deplored the cutback in exchanges with the United States — although the US National Academy of Sciences had asked for talks to continue on international security and arms control. Contacts with the United Kingdom, on the other hand, had, he said, "become more active".

Vera Rich

University research funding

Line of defence

Washington

US universities and colleges continue to benefit from Washington's current enthusiasm for increasing its expenditure on military technologies and related national security needs. A subcommittee of the House of Representatives has now approved a provisional bill, said to have wide support within the Administration, which would allocate up to \$100 million a year to improve facilities in higher education institutions in technical subjects important for security interests.

The explicit purpose of the legislation is to increase the ability of the United States to meet both its immediate defence needs and the technological demands that would be created by a prolonged war.

The legislation is supported by the Department of Defense, the Department

"HERE'S YOUR GRANT—THIRTY THOUSAND PIECES OF SILVER."



of Commerce and congressmen representing areas of the United States which are suffering hardest from the current recession.

Furthermore, according to a report published last month by the Defense Science Board (DSB) on the state of the nation's research universities and their ability to absorb increased military research funds, many universities which had rejected close involvement with the Pentagon in the late 1960s and early 1970s are now experiencing a change of heart.

The DSB study has been used by the Department of Defense to justify the major increase in university research funding which it is asking Congress to support. In his budget request last month, President Reagan proposed a 30.5 per cent increase in university basic research funded by the Army, to a total of \$92.1 million; a 36.9 per cent increase in such research funded by the Air Force, to reach \$89.7 million; and a 17.2 per cent increase in Navy funding. Basic research supported by the Defense Advanced Research Projects Agency would rise to \$8.4 million, a 21.7 per cent increase. Many of the extra funds would go explicitly to raise the level of effort in the mathematical sciences, electronics, chemistry and engineering conversion.

The details of the Defense Department's budget request reflect some of the science board's conclusions and some of the universities' concern. In response to complaints about the declining state of the universities' research facilities, the department has announced its intention to allocate \$150 million over a five-year period to improving research instrumentation. Each of the three major services will put aside \$10 million a year for this purpose.

The department is also attempting to overcome the increasing difficulty of universities in retaining top graduates in fields such as engineering by setting up a new graduate fellowship programme. Stipends will be \$12,000 a year — about 50 per cent higher than comparable awards made by the National Science Foundation — and the services will, in addition, provide affiliated universities with \$8,000 a year to cover tuition and fees.

The initial plans are to award about 125 graduate fellowships in the fiscal year 1983, which starts on 1 October 1982, costing the Department of Defense about \$2.5 million. Dr Jack Crowley of the Association of American Universities has said that the association warmly supports both the proposed funding for research facilities and the new fellowship programme.

If Congress does grant the extra funds there is unlikely to be any shortage of applications for the additional grants. With support for basic research from other agencies by falling in real terms and with industry unlikely to be able to make up much of the shortfall, the universities are looking to the Department of Defense as the only significant alternative source of support.

David Dickson

Oceanographic research

In the black

At a time when many groups are hard pressed to support large projects, a team from the UK Institute of Oceanographic Sciences (IOS) has just returned from a successful six-month cruise on which it earned more than £250,000 in foreign currency. IOS, now run by the Natural Environment Research Council, has been able to share the cost of its unique side-scan sonar project GLORIA (Geological Long-Range Inclined ASDIC) by combining its interests with those of several research groups abroad. With the ability to scan a swathe of the ocean floor almost ten times wider than conventional deep-tow sonars, the equipment is quicker and cheaper than other survey techniques, and many groups overseas are apparently eager to use it.

Research on this multi-project charter — the first long voyage to use GLORIA — was strongly oriented towards commercial applications, principally oil exploration and radioactive waste disposal. The UK Department of Energy commissioned the first part of the voyage, with an extensive survey off Newfoundland aimed at completing a reconstruction of the North Atlantic during the Cretaceous to help



GLORIA aboard the trawler Farnella

assess the oil potential of Britain's continental shelf. The other British project, from the Department of the Environment, was an investigation of the possible disposal of high-level radioactive waste under the sea bed. Searching for specific sites is not the primary objective at present; the department is more interested in using GLORIA's ability to distinguish fine textural changes in the ocean floor to study the varying types of sea-bed environment.

One of the most exciting finds came from an area known as the Amazon Cone off the Brazilian coast. This is a huge pile of sediments washed out by the Amazon into the Atlantic, and is thought to be a potentially important source of oil. Submarine current meanders were for the first time observed to cut into the cone, giving new clues about processes of sediment deposition that may be relevant to the formation of oil-bearing structures. There has been great interest in this finding from petroleum geologists and from the Brazilian oil company Petrobras, which paid for this leg of the cruise jointly with

the Lamont-Doherty Geological Observatory of New York's Columbia University.

During the last leg of the voyage, on the way home from Miami, GLORIA was tested as a means of detecting deposits of manganese nodules, which at present can be located only by trial and error sampling. There seems to be a chance that GLORIA's sensitivity to slight variations in the reflectivity and roughness of the bed may offer a more direct means of location. United States scientists are said to have been impressed with the equipment.

David Millar

Soviet shake up

A regional earthquake monitoring centre has been opened in Dushanbe, in the Soviet republic of Tadjikistan, as a first step towards a seismic early warning system for the whole Soviet Union. Tadjikistan suffers several thousand earth tremors a year, and seismic activity is also very high in the neighbouring republics of Turkmenia, Uzbekistan, Kirghizia and Kazakhstan.

The new centre will combine the resources of existing seismological centres of the Academies of Sciences of the various republics in the area, including a seismic testing ground covering 30,000 square kilometres in the northern Tien-Shan. The facilities of this site include wells 3,000 kilometres deep for studying subterranean waters — in particular the radon content of such waters which has been observed to increase shortly before a tremor. Field stations at the site also monitor the Earth's magnetic field and telluric currents, while microdisplacements of individual crustal blocks are measured with lasers. Biological teams monitor the behaviour of snakes, ants and other fauna said to behave abnormally during the build-up to an earthquake.

In 1980–81 Soviet seismologists say they successfully predicted the occurrence of 13 earthquakes, including force 4–5 tremors in Dushanbe. The seismologists seem very open-minded to any suggestions of possible early-warning signals. Recent investigations have included the relationship of the amount of mercury vapour emitted from the Earth's crust, or present in subterranean well water, and an assessment of local folk lore which has it that a seismic shock is frequently preceded by overcast skies and dust-storms.

The heart of the Dushanbe centre will be a computerized seismological data bank, linking existing centres, to be augmented by links to new mobile seismic monitoring stations scheduled to go into production in the near future.

Vera Rich

CORRESPONDENCE

The return of malaria to India

SIR — In their article on malaria resurgence, Chapin and Wasserstrom¹ claim that the agronomic use of DDT led to the development of DDT resistance in mosquitoes and hence to a resurgence of malaria in India. However, the main cause of the rapid increase in the consumption of insecticides in India since 1970 has been their greater public health use in the wake of malaria resurgence², and agricultural use of DDT has remained fairly constant during the past several years (see table). So the correlation between increase in DDT use and rice production (Fig. 1 of Chapin and Wasserstrom) is superficial and does not imply a cause-effect relationship. Similarly, noting static cotton production and increasing total DDT consumption, Chapin and Wasserstrom are incorrect to conclude that "more and more DDT must be sprayed simply to maintain a fixed yield".

The public health use of DDT in India has always exceeded its agronomic use, so although the agricultural use of insecticides has contributed to the resistance problem, blame must be shared by both agricultural and public health applications. And moreover, as multiple resistance in Indian mosquitoes to DDT, BHC and malathion is not cross-extending³, correlation between the use of a single insecticide (DDT) and malaria cases (Fig. 2 of Chapin and Wasserstrom) can only be spurious.

When seeking reasons for the present problems for malaria control in India, administrative lapses are a factor which must be considered. Under the World Health Organization (WHO), moves in the 1950s to reduce *Plasmodium* transmission by attacking female mosquitoes via insecticide spraying, and to deplete the *Plasmodium* reservoir in human beings by treating patients with antimalarial drugs, were successful. However, this success was nullified because mosquito populations in the forests of Assam, Madhya Pradesh and Orissa were not taken into account. These populations were not affected by spraying of dwellings, so remained infective and gradually reintroduced *Plasmodium* to the now resistant mosquito population. Here it should be noted that Chapin and Wasserstrom use "resistant mosquito populations" and "malaria" as synonyms, when in fact resistant mosquito populations existed harmlessly in India for a decade becoming a hazard only in the presence of *Plasmodium*.

In attempting to find ways of overcoming India's malaria resurgence, an example to learn from is that of California, where use of physical, cultural and biological techniques as adjuncts to chemical control has both prevented malaria resurgence and reduced insecticide use to about a third of the levels used a decade ago⁴. Such techniques have substantially cut the costs of mosquito control in California, and if they were adopted in developing countries, savings could be used to replace DDT with the alternatives which have not been used before because of the high costs.

Although what appears clear with hindsight may in fact have been difficult to foresee, it seems fair to say that WHO was slow to recognize the limitations of residual

Insecticide consumption in India, 1973–1979				
Insecticide	1973–74	1975–76	1977–78	1978–79
<i>Public health</i>				
BHC	5.0	6.7	12.0	17.0
DDT	7.0	7.2	6.0	10.2
Malathion	0.5	1.0	1.1	5.5
<i>Agriculture</i>				
BHC	15.4	18.6	16.6	20.0
DDT	2.9	2.9	2.5	3.0
Malathion	0.4	0.6	1.0	1.5
Carbaryl	2.5	3.5	2.0	3.0

Values are in millions of kilograms. The figures for 1978–79 are for projected consumption.

insecticides. WHO placed prime reliance on chemical control and gave low priority to non-chemical control measures so that when malaria control began to deteriorate, no proven alternatives were available to supplement insecticides. A review of literature on mosquito pest management strategies⁵ reveals that WHO's efforts in this line started in the 1970s, by which time mosquito resistance to insecticides had become rampant. Though it is difficult to differentiate between the residues from insecticide use in agriculture and public health, preliminary studies^{6,7} indicate that DDT and BHC use in malaria control may be responsible for the serious pollution by these components in malarious developing countries⁷⁻⁹. This side effect of malaria control with persistent insecticides has also received limited recognition from WHO.

PARM PAL SINGH

Department of Entomology,
Punjab Agricultural University,
Ludhiana, India

- Chapin, G. & Wasserstrom, R. *Nature* **293**, 181–185 (1981).
- Report of Working Group on Pesticide Industry for the Plan 1978–79 to 1983–84 *Pestic. Inf.* **4**, 31–36 (1978).
- Busvine, J. R. *Insects and Hygiene*, 123–124 (Chapman and Hall, London, 1980).
- Kendrick, J. B. Jr., *Calif. Agric.* **34**, 2 (1980).
- Metcalf, R. L., in *Introduction to Insect Pest Management* (eds Metcalf, R. L. & Luckman, W. H.) 529–564 (Wiley, New York, 1975).
- Kapoor, S. K., Chawla, R. P. & Kalra, R. L. *J. Envir. Sci. Hlth B15*, 545–557 (1980).
- Farvar, M. T. in *The Use of Biological Specimens for the Assessment of Human Exposure to Environmental Pollutants* (eds Berlin, A., Wolff, A. H. & Hasegawa, Y.) 155–163 (Nijhoff, The Hague, 1979).
- Kalra, R. L. & Chawla, R. P. in *Indian Pesticide Industry Facts and Figures* (ed. David, B. V.) 251–285 (Vishwas, Bombay, 1981).
- Al-Shahristani, H. & Mahmood, N. A. *J. Iraqi chem. Soc.* **2**, 69–76 (1977).

Helical liposomes

SIR — Lin *et al.*¹ have recently described the "induction of helical liposomes by Ca^{2+} -mediated intermembrane binding", and the two types of helical liposomes discovered by Lin *et al.* were displayed on the cover of

Nature of 11 March. The authors acknowledged the independent observation of tubular and helical liposomes by another contemporary scientist. But it may be of further interest to your readers that two types of helical liposomes were in fact described some 78 years ago² (see figure).

The book by Lehmann² and a more recent review³ may serve as guides to the extensive studies made after 1854⁴ on tubular and

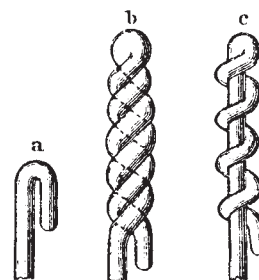


Fig. 477.

helical myelin figures (*zopfartige Flechtwerke*). Lehmann² also offers a mechanism for the formation of the two types of helical liposomes.

H. SANDERMANN

Institut für Biologie II,
Universität Freiburg,
Freiburg i. Br., FRG

- Lin, K.-C., Weis, R.M. and McConnell, H.M. *Nature* **296**, 164–165 (1982).
- Lehmann, O. (*Flüssige Kristalle, sowie Plastizität von Kristallen im Allgemeinen, Molekularumlagerungen und Aggregatzustandsänderungen*) (Engelmann, Leipzig, 1904).
- Kelker, H. *Molec. Cryst. Liquid Cryst.* **21**, 1–48 (1973).
- Virchow, R. *Arch. path. Anat. Physiol. klin. Med.* **6**, 562–572 (1854).

French freedom

SIR — With reference to your important issue, *Science in France* (25 March), may I highlight one crucial advantage that French scientists enjoy over their British counterparts? Applicants for a research grant, which may be rejected, have the statutory right of access, under France's freedom of information laws, to the case-file concerning their application, including referees' reports and other attendant reasons behind the decision, which they may then openly contest and indeed reverse¹.

British scientists who feel that they order this matter better in France, compared, that is, with the closed and secretive system of administration in Britain, would do well to make strong representations to their MPs and trade unions^{2,3}.

STANLEY ALDERSON

Cambridge, UK

- New Scientist*, **83**, 64 (1979).
- Alderson, S. *Nature* **278**, 206 (1979).
- Alderson, S. *Times Higher Education Supplement*, 7 December 1979.

NEWS AND VIEWS

Geochemical light on mantle convection

from Donald L. Turcotte

MANTLE convection is generally accepted to be the driving mechanism for continental drift and plate tectonics, but controversy persists about the scale on which the process takes place. Does convection involve the whole mantle, down to the boundary with the core? Or, as Allègre¹ has recently suggested on geochemical grounds, are only the upper layers of the mantle involved? Is there even, as Houseman and McKenzie² suggest, a second scale of mantle convection?

The controversy centres on the seismic discontinuity at a depth of 650 km, where seismic velocities increase by about 10 per cent. The increase can be attributed either to a solid-solid phase change or to a change of composition, perhaps of iron content. The absence of seismicity below this depth is taken as evidence that the convection associated with plate tectonics is limited to the upper mantle.

Geochemical observations lend further support to the notion of layered mantle convection. Systematic studies³⁻⁵ of isotopic ratios in the samarium-neodymium system have shown that the upper mantle and the continental crust are complementary geochemical reservoirs. The upper mantle reservoir is systematically depleted in those elements, the rare earths for example, which are incompatible with the mantle material and the continental crust is correspondingly enriched. The mass balance of the Sm-Nd system suggests that 30±10 per cent of the upper mantle has been involved, which is consistent with an upper mantle convection system limited to depths of less than 650 km.

A possible reservoir system is illustrated in the figure. Mid-ocean ridges, which randomly tap the upper mantle reservoir, usually yield mid-ocean ridge basalts which are similar in both major and trace element chemistry, showing that the upper mantle

reservoir is well stirred. Volcanism associated with the subduction zones systematically transfers incompatible elements into the continental crust. The remainder of the subducted lithosphere, including the enriched oceanic crust and depleted mantle, must then be mixed into the upper mantle reservoir in order to provide the nearly uniform and depleted basalts found at mid-ocean ridges.

This simple two-reservoir model does not, however, explain all the geochemical observations. Ocean island basalts (for example, Hawaii) are not systematically depleted but exhibit a range of geochemical signatures, from depleted to enriched. One explanation is that ocean island basalts result from partial melting in a plume rising from an undepleted lower mantle reservoir⁶, in which case the range of compositions is attributed to mixing between materials of the depleted upper mantle and the undepleted lower mantle. This model does not, however, explain all enrichments of ocean island basalts, whence the proposal⁷⁻⁹ that the source of the ocean island basalts is subducted oceanic crust, enriched with continental trace elements by seawater alteration and by the entrainment of sea-floor sediments¹⁰, that has not been fully mixed into the mantle.

In passing, the two-reservoir model must

also be amended to take into account the discordance between geochemical studies of the Sm-Nd system on the one hand and the strontium-rubidium and uranium-lead systems on the other. Fractionation within the crust seems to have depleted its upper layers of lead and strontium, suggesting a differentiation of the crustal reservoir into two. Whether the 'missing' lead is in the lower crustal reservoir or (as sulphide) in the molten core is disputed.

While these geochemical studies favour a layered mantle convection system, there are some difficulties. Thus Jordan and Creager¹¹ have suggested that travel time anomalies indicate the penetration of the descending surface plate to depths of at least 900 km, which argues against stratified convection.

If indeed the mantle is chemically stratified at 650 km or thereabouts, independent convection systems would be expected in the upper and lower mantle. If the observed discontinuity marks a phase change, its thermal coupling between the two systems will depend on the thermodynamic properties of the phase change. High-pressure measurements¹² using diamond cells with laser heating indicate that the principal upper mantle mineral olivine transforms to basic oxides (magnesiowustite plus perovskite) at conditions reasonably close to those at a

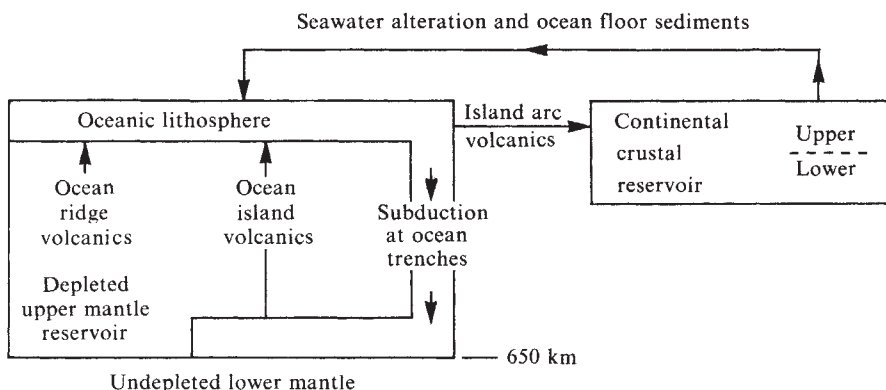


Illustration of the mantle reservoirs.

Donald L. Turcotte is Chairman of the Department of Geological Sciences at Cornell University, Ithaca, New York 14853.

depth of 650 km, but not enough is known about this transformation to decide whether it would assist or block convection. Moreover, Liu¹³ argues that there is a significant discrepancy between predicted and required conditions for the olivine transformation, and favours chemical stratification.

Further objections have been suggested by the laboratory and numerical studies of stratified convection carried out by Richter and McKenzie¹⁴, who predict a double thermal boundary layer at the interface between the convecting layers with a significant increase of temperature across it. But this increase of temperature would imply¹⁵ a large (several orders of magnitude) decrease in viscosity which is not consistent with observation of postglacial rebound. This conclusion is a serious difficulty for the concept of stratified convection which is unlikely to be recovered by a consideration of dependence of the viscosity or rheological differences between the convecting layers.

That there may be a second and smaller scale of mantle convection is suggested by several observations. One is intraplate volcanism, such as that at Hawaii and other oceanic islands. Another is the requirement that the upper mantle reservoir be well stirred or mixed in order to have a nearly uniform composition. Richter and Parsons¹⁶ suggested that instabilities within the lithosphere may be responsible, but Yuen, Peltier and Schubert¹⁷ have questioned the validity of the numerical

studies so far made of this process² on the grounds that they do not allow for the strong temperature dependence of the mantle viscosity. They go on to suggest that the hot, low-viscosity lower boundary layers are more likely to be unstable than the lithosphere. And, such unstable lower boundary layers could help them to explain the ocean island volcanism.

It is, of course, too soon to know how these questions will be decided. What is, however, already plain is that geochemical methods will throw light on the mechanisms of mantle convection, which are not merely the driving forces of continental drift but powerful agents in the evolution of the planet. □

1. Allègre, C.J. *Tectonophysics* **81**, 109 (1982).
2. Houseman, G. & McKenzie, D.P. *Geophys. J. R. astr. Soc.* **68**, 133 (1982).
3. Jacobsen, S.B. & Wasserburg, G.J. *J. geophys. Res.* **84**, 7411 (1979).
4. O'Nions, R.K., Evenson, J.N. & Hamilton, P.S. *J. geophys. Res.* **84**, 6091 (1979).
5. Allegre, C.J., Othman, D., Polve, M. & Richard, P. *Phys. Earth planet. Inter.* **19**, 293 (1979).
6. Schilling, J.G. *Nature* **242**, 565 (1973).
7. Chase, C.G. *Earth planet. Sci. Lett.* **52**, 277 (1981).
8. Hofmann, A.W. & White, W.M. *EOS* **62**, 424 (1981).
9. White, W.M., Vosage, H. & Hofmann, A.W. *EOS* **62**, 424 (1981).
10. Karig, D.E. & Kay, R.W. *Phil. Trans. R. Soc. A* **301**, 233 (1981).
11. Jordan, T.H. & Creager, K.C. *EOS* **62**, 1074 (1981).
12. Yagi, T., Bell, P.M. & Mao, H.K. *Carnegie Inst. Wash. Yb.* **78**, 614 (1979).
13. Liu, L. *Earth planet. Sci. Lett.* **42**, 202 (1979).
14. Richter, F.M. & McKenzie, D.P. *J. geophys. Res.* **86**, 6133 (1981).
15. Schubert, G. & Spohn, T. *Geophys. Res. Lett.* **8**, 951 (1981).
16. Richter, F.M. & Parsons, B. *J. geophys. Res.* **80**, 2529 (1975).
17. Yuen, D.A., Peltier, W.R. & Schubert, G. *Geophys. J. R. astr. Soc.* **65**, 171 (1981).

studies on animals?

Following the original report of transplacental carcinogenesis in pregnant female mice exposed to urethane¹¹, there have been a number of studies on tumour incidence in offspring of pregnant animals exposed to chemical mutagenic carcinogens. In some cases, inbred lines of mice¹² or rats¹³ have been used and brother-sister matings followed through to the F₂ or F₃ generations. All reports show that treatments that result in tumours in exposed individuals and their immediate descendants also result in an excess of tumours in F₂ and F₃ animals. That the enhanced tumour incidence in later generations is not due to transmission of the carcinogen through mother's milk or excreta was ruled out through the use of rapidly metabolised and inactivated mutagens such as *N*-nitrosomethylurea (NMU)¹³ or *N*-nitrosoethylurea (ENU)¹⁴. Moreover, in a recent study, Tomatis *et al.*¹⁵ have shown a similar increased incidence of tumours in F₁ offspring of male rats treated with ENU some weeks before mating, so that a mutagen-induced enhancement of tumour incidence can also be transmitted through the male germ line. The story is taken further by the large-scale study reported by Nomura in this week's *Nature* in which tumour incidence in offspring of male and female mice exposed to X rays or urethane has been followed over a number of generations.

On page 575, Nomura summarizes a study extending over ten years, and involving some 15,000 mice, and reports that exposure of late-stage male or female germ cells to X rays gives a linear increase in dominant lethals with dose, but no increase when animals are exposed to the mutagen urethane; however, parental exposure to both agents resulted in congenital abnormalities in F₁ embryos. Of special interest is the finding that both X-irradiated and urethane-treated male or female mice produced progeny with a much increased tumour incidence, with a clear dose-response relationship for offspring derived from irradiated late-stage germ cells. At the highest X-ray dose of 500 rads, the tumour incidence in the offspring of irradiated animals (~30 per cent) was some six times that in the controls (5 per cent). The spectrum of tumours in controls and in offspring of treated parents was similar, with some 90 per cent being non-malignant papillary adenomas of the lung and the remainder various malignant tumours of other organs. Matings continued down to the F₃ generation clearly demonstrated heritability of the induced high tumour incidence and the pattern of inheritance is that of a dominant trait with 40 per cent penetrance. Although the X-irradiation of males resulted in dominant lethals and cytologically demonstrable translocations, no such effects were observed in animals treated with urethane, and Nomura argues that the mutations that result in excess tumours are probably not gross

Parental mutagenesis and familial cancer

from H.J. Evans

PEOPLE can inherit mutations which predispose their offspring to the early development of specific cancers, and which are loosely referred to as 'cancer genes'¹; and mutagen-induced chromosomal rearrangements, or more subtle mutations, may initiate malignant transformation in exposed somatic cells². Does this mean that exposure to mutagenic carcinogens results not only in an increased incidence of cancer in exposed individuals, but also in their descendants?

It is recognized that abdominal exposure of pregnant women to X-rays may result in an increased incidence of leukaemia in the resultant offspring⁴; that maternal exposure to synthetic oestrogens during pregnancy can result in the subsequent development of vaginal adenocarcinoma in female children⁵; and that exposure of children to asbestos dust brought home from the work place by an occupationally exposed parent is responsible for the later developmental of mesothelioma in exposed offspring⁶. More recently, an excess of

brain tumours has been noted in children of adults occupationally exposed to chemical solvents⁷. But in all these studies, the excess cancers are the result of the exposure of the offspring themselves and are not a consequence of gonadal exposure of parents and transmission of a cancer-predisposing genotype to the offspring. Indeed, the evidence in man for transmission of any kind of induced germ cell mutations to offspring by parents exposed to chemical mutagens is almost non-existent, with some limited evidence from studies of individuals exposed to vinyl chloride⁸ or to anaesthetic gases^{9,10}. There is no information available on any possible germ cell transmission of an acquired cancer risk from parent to offspring. What evidence do we have from experimental

H.J. Evans is Director of the MRC Clinical and Population Cytogenetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XJ.

chromosomal rearrangements.

The rate of X-ray induction of mutations that predispose to tumour development turns out to be around ten times greater than the mutations that, for example, result in skeletal changes in the mouse and Nomura suggests that there may be a large number of sites in the genome where mutation may predispose to tumour development. Other explanations are possible, but whatever mechanism is involved it is clear that mutagen exposure results in transmitted heritable changes that considerably enhance tumour incidence in the offspring. The demonstration of this effect required parental exposure to fairly high doses of mutagens and the use of inbred lines and we may ask whether these results are of significance in relation to human exposure to mutagens. There is no reason to believe that effects similar to those in the mouse may not occur in man, although they would be less easily demonstrated. In this context, two human populations of interest would be the offspring of the survivors of Hiroshima and Nagasaki who were exposed to high radiation doses, and who showed an enhanced cancer incidence and chromosome damage in their blood cells¹⁶, and the

offspring of cigarette smokers who must comprise one of the largest groups of individuals exposed to a chemical mutagen at dose levels sufficient to produce significant chromosome damage in their blood lymphocytes¹⁷. □

1. Knudson, A.G. in *Genes, Chromosomes, and Neoplasia* (eds Arrighi, F.E., Roa, P.N. & Stubblefield) 453 (Raven, New York, 1981).
2. Evans, H.J. in *Proceedings of Symposium on Host Factors in Human Carcinogenesis* (IARC, Lyon, in the press).
3. Bittner, J.J. *Science* **95**, 462 (1942).
4. McMahon, B. & Hutchinson, G.B. *Acta Un. int. Cancer* **20**, 1172 (1964).
5. Herbst, A.-L., Kurman, R.J., Suilley, R.E. & Poskanzer, D.C. *New Engl. J. Med.* **287**, 1259 (1972).
6. Wasserman, M., Wasserman, D., Steinitz, R., Katz, L. & Lamesch, C. in *Biological Effects of Mineral Fibres* Vol. 1 (ed. Wagner, J.C.) 253 (IARC Scientific Publication No.30, 1980).
7. Peters, J.M., Preston-Martin, S. & Yu, M.C. *Science* **213**, 235 (1981).
8. Infante, P.F., Wagoner, J.F., McMichael, A.J., Waxweiler, R.J. & Falk, H. *Lancet* **i**, 734 (1976).
9. Cohen, E.N. *et al. Anesthesiology* **41**, 321 (1974).
10. Knill-Jones, R.P., Newman, B.J. & Spence, A.A. *Lancet* **ii**, 807 (1975).
11. Larsen, C.D. *J. natn. Cancer Inst.* **8**, 63 (1947).
12. Tomatis, L. & Goodall, C.M. *Int. J. Cancer* **4**, 219 (1969).
13. Tomatis, L., Hilfrich, J. & Turosov, V. *Int. J. Cancer* **15**, 385 (1975).
14. Tomatis, L., Hilfrich, J. & Turosov, V. *Int. J. Cancer* **19**, 240 (1977).
15. Tomatis, L., Cabral, J.R.P., Likhachev, A.-J. & Pomonarkov, V. *Cancer* **28**, 475 (1981).
16. Awa, A.A. *Radiat. Res. Suppl.* (1975).
17. Evans H.J. in *Progress in Mutation Research*, Vol.2 (ed. Kappas, A.) 111 (Elsevier, Amsterdam, 1981).

With increasing molecular size, however, new problems arise. First, there is the inability to resolve the resonances; increasing molecular weight increases not only the number of resonances but also their peak width. With hundreds or even thousands of proton resonances spread over a frequency range of only a few kilohertz, considerable overlap is inevitable. Even with resonances that can be resolved, an even greater problem is presented by their assignment. Not all the factors affecting chemical shift can be analysed exactly, so that a peak at a given frequency can only be definitely assigned to its corresponding nucleus by direct experiment. This often means chemical modification of the molecules, or comparison of almost-identical proteins, and the labour required for even one or two positive assignments in the spectrum of a protein can be enormous. Once assigned, however, the chemical shift, splitting and other features of the resonance can provide sufficient information about its local chemical environment, including dynamic behaviour, to make the effort well worthwhile.

The work presented by Wüthrich *et al.* represents the culmination of several years of development, in their laboratory and others (notably that of Ernst), of new NMR techniques which not only provide solutions to the problems of resolution and assignment, but further point to a scheme for the determination of protein structure in three dimensions directly from NMR experiments. The methods are those of two-dimensional NMR, in which the sample is subjected to a sequence of radio-frequency pulses of the general form 'pulse-evolution time t_1 - pulse - data acquisition time t_2 '. The data are thus collected as a function of t_2 with t_1 as a parameter, and if a large number of experiments are performed for different values of t_1 the result is a matrix of data points, each of which is a function of both t_1 and t_2 . This matrix is Fourier-

New dimensions in protein NMR

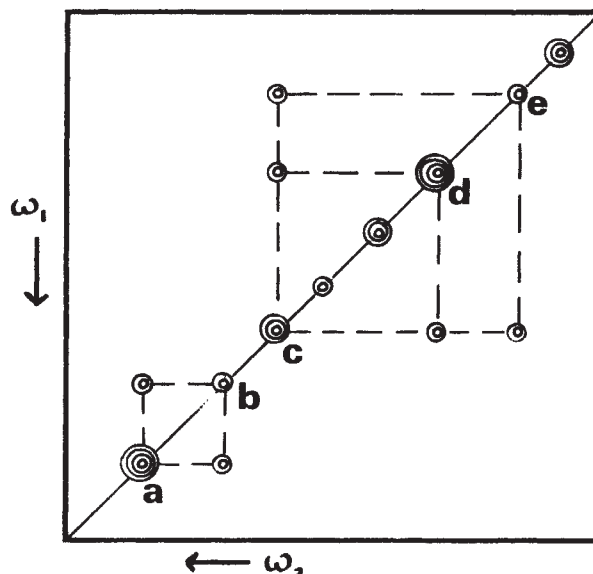
from H. W. E. Rattle

NUCLEAR magnetic resonance (NMR) has been much in the news of late; the obvious biochemical and medical applications of such techniques as topical magnetic resonance (TMR) and NMR scanning have aroused a great deal of interest. Arguably, though, an even greater revolution has been going on in the application of NMR to the determination of the molecular structure of proteins, and a recent sequence of papers¹⁻⁴ from Wüthrich and his colleagues shows clearly that it may soon be possible to derive structural information from peptides and the smaller proteins in solution at the same levels of resolution as we have become accustomed to seeing from X-ray crystallography.

When the first protein samples were put into 40 and 60 MHz NMR spectrometers in the mid 1960s, they produced a spectrum of three or four broad featureless humps which made it clear that much work would be needed before useful information could be obtained. The advance of high-field magnet technology produced great improvements, as did the rapid data acquisition made possible by the Fourier transform method. The nuclear spins of the sample are perturbed by the application

of a radio-frequency pulse, and their subsequent collective behaviour monitored as a function of time. Fourier transformation of the resulting amplitude-time function produces the NMR spectrum with each resonant nucleus in the molecule giving rise to a peak at a position on a frequency scale (the chemical shift) which depends on its precise environment within the molecule.

Schematic representation of a two-dimensional COSY or NOESY contour-map spectrum. The peaks of the conventional proton magnetic resonance spectrum appear along the diagonal, and cross-peaks reveal the connectivity between resonances a and b, and between c, d and e. (Adapted from ref.5.)



H.W.E. Rattle is in the Biophysics Group at Portsmouth Polytechnic, St Michael's Building, White Swan Road, Portsmouth PO1 2DT, UK.

transformed twice to yield a two-dimensional spectrum in which peaks appear as functions of two frequencies ω_1 and ω_2 . In the presentation favoured by Wüthrich *et al.*, the spectrum is plotted as a square contour map with the peaks of the conventional spectrum appearing along a diagonal. Clearly more data can be presented in two dimensions than in one, but the real advantage goes far beyond this. By the selection of pulse sequences and appropriate values for the evolution time, each pair of nuclei which interacts within the molecule is made to reveal its 'connectivity' by the appearance of an off-diagonal peak (a cross-peak) on the spectrum. The connectivities may be through-bond spin-spin couplings (J-couplings), in which case the technique is termed two dimensional-correlated spectroscopy (COSY), or across-space connectivities produced at ranges of up to a few Ångström units by the nuclear Overhauser effect (NOESY) (see the figure). Since nuclear Overhauser effects vary with distance, it should eventually be possible to measure short internuclear distances between nuclei not directly linked by bonds; for which at present only approximate values are obtainable.

The general scheme for using COSY and NOESY proton spectra is then as follows: spectra are obtained from a peptide of known sequence in aqueous (H_2O) solution, and at least one backbone α -CH or N-H resonance assigned by conventional means. Spin coupling of this resonance to those of adjacent residues along the chain produces cross-peaks, thus permitting assignment of peaks from the neighbouring residues. These in turn have cross-peaks with the next residues, and so the whole protein backbone may be assigned sequentially. Across-space NOE connectivities obtained from the NOESY spectrum add to and confirm the assignments; detailed analysis of known protein structures show that between 70 and 95 per cent of close approaches are between contiguous residues. Virtually complete assignments of every proton in the molecules of basic pancreatic trypsin inhibitor (58 residues) and membrane-bound glucagon (29 residues) are presented as proof of the efficacy of the method. Given such full spectral assignments, an even more interesting prospect appears, since NOESY spectra can give at least semi-quantitative values for the distances between the α -CH proton of the i th residue and the backbone NH proton of the $(i+1)$ th, between backbone NH protons of adjacent residues, and between β -CH or β -CH₂ of residue i and NH of residue $(i+1)$. The conformation of a residue is as completely specified by these distances as it is by the more conventional Ramachandran bond angles ψ , ϕ and χ . Given enough NOESY measurements of such distances, even if only approximate, and using the other known constraints on the folding of the protein, it should be possible

with appropriate computer methods to arrive at a fairly complete determination of protein conformations in non-crystalline environments, at least for proteins smaller than 70–100 residues. Of course COSY, NOESY and other two-dimensional NMR methods may also be used in conjunction with the many valuable conventional methods, such as paramagnetic probes, relaxation measurements, biochemical modifications and others which are already in regular use. The instrumental requirements for two-dimensional NMR are considerable: high spectrometer

frequencies (300–500 MHz) coupled with computing facilities capable of handling a two-dimensional spectrum which may be the equivalent of several hundred conventional spectra. The advances made possible, however, are very considerable. □

1. Wüthrich, K., Wider, G., Wagner, G. & Braun, W. *J. molec. Biol.* **155**, 311 (1982).
2. Billeter, M., Braun, W. & Wüthrich, K. *J. molec. Biol.* **155**, (1982).
3. Wagner, G. & Wüthrich, K. *J. molec. Biol.* **155**, 347 (1982).
4. Wider, G., Lee, K. H. & Wüthrich, K. *J. molec. Biol.* **155**, 367 (1982).
5. Nagayama, K., Kumar, A., Wüthrich, K. & Ernst, R. R. *J. Mag. Res.* **40**, 321 (1980).

The big bang — 'free lunch' at the Royal Society

from Craig Hogan

COSMOLOGISTS seem to have spent a great deal of time over the years talking about the same things over and over again: element abundances, background radiation, galaxy formation and the like. In so doing they tend to oscillate between congratulating themselves on how much they think they understand and despairing that they will never know anything for certain. Both of these moods were much in evidence at a discussion meeting held recently at the Royal Society, supposedly on "The Big Bang and Element Creation" but actually covering a wide range of subjects intended to span the whole range of observational cosmology and a good deal of theoretical cosmology.

As expected, the most genuinely new results came from groups measuring anisotropy in the microwave background radiation. D.T. Wilkinson (University of Sussex), whose experiment is capable of measuring large-scale (90°) angular fluctuations in radiation temperature with a system noise of only 0.1 mK (compared with the total 3K background temperature), described a possible observation of a cosmological quadrupole moment with amplitude $\delta T/T \approx 10^{-4}$ which, if true, would be an observation of historic significance, because such measurements observe structural properties of the Universe at very early times, possibly even the big bang itself. Unfortunately, as Wilkinson noted, such observations are beset with numerous systematic difficulties expected to be only slightly fainter than the observed signal, and this became obvious when he showed a sky map made with a new low-noise maser detector in which the Galaxy appeared as a bright band that was much more conspicuous than any trend in background brightness. R. Fabbri (University of Florence) described results from the

Florence group and cited tentative evidence that cosmological fluctuations at 6° angular scale are smaller than the quadrupole ($\delta T/T = 3 \times 10^{-5}$) and decreasing with angle — a result as exciting as Wilkinson's, but difficult to reconcile with it in any simple picture. On the other hand, Wilkinson showed very clearly that the dipole moment ('the great cosine in the sky' due to the Earth's motion) is now firmly established, even if minor differences still exist between various groups.

In the same tone, P.L. Richards (University of California, Berkeley) showed that the spectrum of the background radiation is indeed very nearly blackbody down to submillimetre wavelengths, even though controversial apparent departures from an exact thermal spectrum may exceed the statistical limits. Woody and Richards' technological achievement may be appreciated by noting that the cosmological background they measure had a brightness only 10^{-5} of an ordinary body at room temperature. New experiments were described which may check the significance of the spectral distortion, which again, if real, gives us significant positive information about the early evolution of the Universe.

Measurements of the cosmic abundance of the elements — to date, the only independent verification that the microwave background is primordial — were reviewed by B. Pagel (Royal Greenwich Observatory). Progress has been made here too, with the best measurements (optical observations of HII regions) of primordial helium abundance by mass giving $Y_p \sim 0.23 \pm 0.01$, a conservative statement being $Y_p < 0.25$. As pointed out by D. Schramm (Enrico-Fermi Institute, University of Chicago), with $Y_p = 0.25$ the hot big bang predicts the observed abundances for deuterium and ^3He with remarkable success.

However, the helium observations now

Craig Hogan is in the Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA.

also place serious constraints on the standard hot big bang model. M. Davis (Center for Astrophysics, Cambridge, Massachusetts) showed evidence from the flow of galaxies in the Virgo supercluster that the total mass density of the Universe is at least one-fifth of the density which would be necessary to keep it from expanding forever — in terms of the conventional cosmological density parameter, $\Omega_{\text{TOTAL}} \gg 0.2$. Most of this mass is dark material whose nature is unknown. If, however, it were made of ordinary matter, the hot big bang predicts that the amount of helium produced would exceed the observed limits. If the standard model of helium synthesis is correct, then $Y_p < 0.25$ implies that the fraction of closure density in baryons is $\Omega_b < 0.1$. Therefore, the standard model already requires that most of the mass in the Universe is some other form of matter — for example, massive neutrinos, or black holes which formed before nucleosynthesis, or something even more exotic. As Schramm pointed out, the standard model would be ruled out altogether if Y_p proved to be less than Pagel's 'best' value of 0.23. In view of this sensitivity, the success of the theory must be regarded as even more remarkable.

W. Sargent (Caltech) pointed out that deuterium may appear in high-resolution spectra of very distant objects — the Lyman- α absorption due to hydrogen clouds in front of high-redshift quasars. Such systems provide direct, if fragmentary, information about the intergalactic gas at redshift of order 2, and have already been used by Sargent and his collaborators as a probe of galaxy formation processes.

The simplest prediction of all was made by A. Guth (MIT), who showed that in some circumstances, a tiny region (or 'bubble') of primeval vacuum could inflate to include the entire observable universe, then decay to fill it with all of the matter and radiation.

The reason such a scheme is possible in the context of modern gauge theories is that the primordial vacuum is allowed to have a large energy density. When such a 'false' vacuum expands, its energy density does not decrease (like ordinary matter) but stays the same, which enables the universe to expand by a very large factor while remaining at the same density. The universe becomes 'normal' when the energy of the vacuum is converted into ordinary matter and radiation (by a decay process which is not well understood, and which is the focus of research), and the vacuum becomes the normal, non-gravitating 'true' vacuum we have today. (We know that the energy density of the present-day vacuum is nearly zero because the 'cosmological constant', which measures its effect on the present-day expansion, is shown by observations of distant galaxies to be nearly zero.)

Such an 'inflationary universe' auto-

matically predicts that $\Omega_{\text{TOTAL}} = 1 \pm \epsilon$ (where $\epsilon \ll 1$), and, because of the uniformity of the vacuum, also predicts that no primordial fluctuations are produced — the early Universe is both very flat and very uniform. It appears that in this kind of scheme, one will have to find an explanation for the observed non-uniformity in terms of processes occurring after the inflation. However, the most interesting of Guth's speculations is that the bubble of vacuum, and hence the entire universe, could arise by quantum fluctuation 'from nothing'. Apparently certain kinds of gauge symmetry might predict not only the structure of matter but

also the structure and existence of the space time in which the matter fields operate.

Nowadays, not only does modern physics give you electrons and atoms, but it also gives you universes. Guth's vision of the Universe as a 'free lunch' will appeal to many. I walked away afterwards (as it turned out, to go to lunch) feeling perplexed as if by a *koan* — if the Universe owes its existence and symmetry to gauge theory, does the gauge theory owe its 'realization', and hence its existence or 'reality', to the Universe it creates? No doubt these speculations will provide cosmologists with plenty of food for thought. $\square$

When is a zeolite not a zeolite?

from Lovat V.C. Rees

In 1972 the Mobil Oil Corporation synthesized and patented a range of very high Si/Al zeolites of which ZSM-5 is the best known example. This zeolite is a catalyst of great importance in the synthesis of petrol from coal (see Thomas *et al.*, this issue of *Nature* p.530) and the Mobil patent is thus of considerable commercial significance. As Thomas *et al.* indicate, the Union Carbide Corporation has more recently synthesized and patented a new, thermally stable, crystalline phase of silica, called silicalite, which is topologically very similar, if not identical, to ZSM-5.

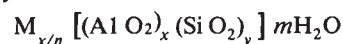
Silicalite, however, always contains some aluminium (Si/Al ratios >200 are usually found) which Union Carbide assert is present in the form of Al_2O_3 impurity. The Mobil patent for ZSM-5 embraces Si/Al ratios of up to 4,000, so the zeolite may contain less aluminium than some samples of the silicalite. If some of the aluminium in silicalite is present as tetrahedral framework aluminium, then a philosophical argument arises as to whether a sample of ZSM-5 (US Patent 3702886, 1972) with very low aluminium content is in any way different from silicalite (US Patent 4061724, 1977). This philosophical argument, which could have considerable financial implications, can be carried a stage further when one asks "When is a zeolite not a zeolite but a defect silica"?

If the aluminium content decreases below one aluminium atom per unit cell, can we still consider the material to be an aluminosilicate with a specific unit cell formula or is it a defective form of crystalline silica? Even if one aluminium per unit cell is present, if this aluminium atom is randomly distributed among all tetrahedral sites in the unit cell is this material

any more a zeolite than material containing less than one aluminium per unit cell?

The paper from Thomas and his colleagues at Cambridge and Guelph goes to the heart of this question. They have shown by ^{27}Al magic-angle spinning NMR (MASNMR) that all of the aluminium atoms in a very 'pure' silicalite (Si/Al $>1,000$) and the less 'pure' commercial silicalite (Si/Al >200) are located in tetrahedral sites. No octahedral aluminium was detected. They also found a small signal in their ^{29}Si MASNMR spectrum of the commercial silicalite due to Si(1Al) which must be present if the aluminium is present in tetrahedral co-ordination. If these results are accepted then we must draw the conclusion that silicalite is essentially indistinguishable from a high Si/Al ratio ZSM-5 and leave the question of whether these materials are

Zeolites are crystalline aluminosilicates which are widely distributed in nature but are also readily synthesized in the laboratory. The general formula for a zeolite may be written as:



where M is a cation of valency n and the Si/Al ratio can range from one to very large values. In many zeolites, the zeolitic water can be readily removed without lattice collapse, leaving a porous crystal with void volumes which can be as large as 50 per cent of the crystal volume. This void volume consists of a network (frequently three-dimensional and interconnected) of channels of molecular dimensions. Zeolites have found wide use, therefore, as molecular sieves and as acid catalysts. When the M cations are replaced by H^+ ions, very active cracking, hydrocracking and isomerization catalysts are produced which are extensively used by the petroleum industry.

Lovat V.C. Rees is Reader in Physical Chemistry at the Imperial College of Science and Technology, London SW7 2AY.

zeolites or not unanswered.

Another novel application of ^{29}Si and ^{27}Al MASNMR, and some further exciting results, are presented in this issue in a second, joint contribution from the above two schools (see p.533). Zeolite Y (synthetic faujasite) as synthesized is an excellent cracking catalyst but is found to be inadequate under plant use when high temperatures are employed in the decoking process. The conversion of zeolite Y to an ultra-stable form was an important development which has led to the production of >200,000 tons of ultra-stable zeolites per annum. Although the

main features of this stabilization process have been indicated in previous publications, the new work presents convincing evidence of the fate of the aluminium atoms removed from tetrahedral framework sites and of the healing of the vacancies so produced by silicon atoms which come, not exclusively from surface or amorphous regions of the crystal, but also from the bulk crystal. A direct probe has been brought to bear, therefore, to monitor the internal reorganization of tetrahedral atoms which converts a catalyst into a commercially viable catalyst.

What next will we find by MASNMR? □

Is phosphatidylinositol really out of the calcium gate?

from Robert H. Michell

SEVEN years ago I suggested that receptor-stimulated breakdown of phosphatidylinositol (PI) might be implicated in a general mechanism whereby activation of many cell-surface receptors brings about a rise in the cytosolic Ca^{2+} concentration in stimulated cells¹. In recent months, two brief reviews, one by J.N. Hawthorne in these columns, have argued that this general hypothesis is no longer tenable^{2,3}. By contrast, a number of other recent reviewers have continued to treat it as a useful working hypothesis, despite emphasizing that it is still unproved⁴⁻⁸. I would like briefly to summarize some of the weaknesses of Hawthorne's recent *News and Views* article³. I feel that these criticisms are sufficiently fundamental to raise serious doubts about the purported disproof of the postulated link between stimulated inositol lipid breakdown and the mobilization of Ca^{2+} in stimulated cells.

Hawthorne considers that the proposed mechanism for Ca^{2+} mobilization is cumbersome, mainly because not all the necessary enzymes are at a single cellular site, and that it is wasteful of energy. But membrane lipids are always in a state of continuous metabolic turnover, and the rapid synthesis to support this continuous turnover occurs mainly at the endoplasmic reticulum. The inositol lipids of the plasma membrane that are involved in cell responsiveness may simply be renewed by intracellular synthesis rather more rapidly than the other phospholipids of the plasma membrane. When considering how much energy this costs, a comparison with adenylate cyclase may be instructive. Synthesis of cyclic AMP expends two high-energy phosphate bonds (plus an unquantified consumption of GTP by the guanyl nucleotide-binding coupling protein) and yields a single molecule of the intracellular

messenger. Breakdown and resynthesis of one molecule of PI consumes three high-energy phosphates (assuming conservation and reutilization of the diacylglycerol moiety). If the initiating event turns out to be phosphodiesterase attack on PI 4,5-bisphosphate ($\text{PI}4,5\text{P}_2$; see below), then this expenditure will rise to five high-energy phosphate bonds. As a result of this breakdown, we can assume the release of some membrane-bound Ca^{2+} and/or the opening of a Ca^{2+} channel. Thus the breakdown of one lipid molecule might allow the mobilization of many Ca^{2+} ions into the cytosol compartment of the cell. The three (or five) phosphate bonds seem likely to be a minor energetic consideration, since most energy expenditure during Ca^{2+} signalling undoubtedly occurs in the maintenance of the large Ca^{2+} gradient at the plasma membrane by continuous pumping of Ca^{2+} out of the cells, rather than in the control of Ca^{2+} re-entry.

As Hawthorne mentions, there are situations in which the admission of Ca^{2+} to cells, either by receptor activation or by using an ionophore, can cause the disappearance of PI, but in no case has the route of breakdown been fully documented.

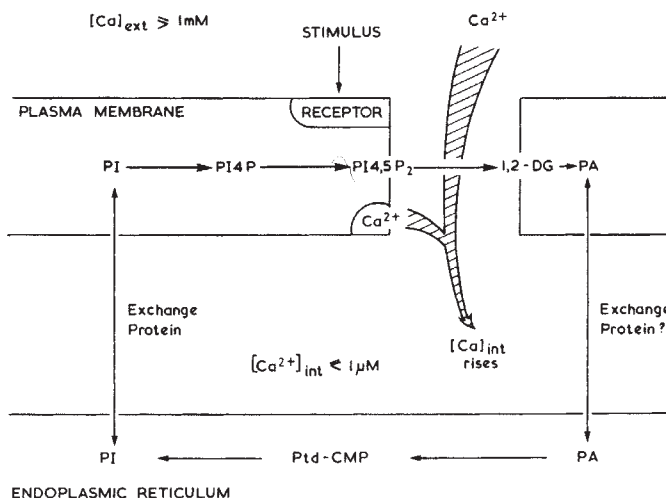
Much of the best evidence comes from platelets, where receptor-coupled disappearance of PI almost certainly involves the activity of a phosphodiesterase (a phospholipase C), whilst Ca^{2+} -mediated PI disappearance is caused by an acylhydrolase (a phospholipase A)^{9,10}. Only the former is of any immediate interest when considering the possible coupling function of receptor-controlled breakdown of inositol lipids. Neither the experiments of Farese¹¹ on pancreas nor of Cockcroft¹² on leukocytes demonstrate the route of Ca^{2+} -dependent PI 'breakdown' in receptor-stimulated cells.

Hawthorne correctly points out that the data on bovine adrenal medulla are not easily reconciled with the inositol lipid breakdown/ Ca^{2+} mobilization hypothesis. Additional information on this system should be illuminating.

It has generally been considered that inositol lipid metabolism is stimulated by activation of muscarinic cholinergic receptors at both pre- and postsynaptic sites, and Hawthorne states this view as a fact. Recent work, however, has shown that in 'synaptosomal' fractions at least a proportion of this response is associated with postsynaptic elements¹³, and I¹⁴ have offered additional arguments to suggest that there may be no presynaptic stimulation of PI metabolism by muscarinic cholinergic stimuli. If it were to prove the case, then the failure of presynaptic muscarinic receptors to mobilize Ca^{2+} would tally well with their failure to stimulate inositol lipid metabolism.

Hawthorne states that "there are so many systems in which the receptor-linked PI loss requires external Ca^{2+} that a universal role of PI in Ca^{2+} mobilization cannot be upheld". This statement is fundamentally flawed, since it fails to recognize that a requirement for extracellular Ca^{2+} does not constitute proof that a change in cytosol Ca^{2+} concentration controls PI breakdown (see ref. 8 for a detailed discussion).

For the past few years it has been generally held that the first lipid to breakdown in stimulated cells is PI rather than one of its phosphorylated derivatives,



A scheme illustrating the possible role of inositol lipids in receptor-controlled Ca^{2+} mobilization in cells. A synthesis based on refs 1 (Fig. 4), 18 (Fig. 2) and 8 (Fig. 3).

Robert H. Michell is Senior Lecturer in the Department of Biochemistry, University of Birmingham, PO Box 363, Birmingham B15 2TT.

PI4P and PI4,5P₂. Breakdown of either one or both derivatives was only reported in conditions likely to raise cytosol Ca²⁺ concentrations greatly, and studies of the enzyme responsible for their Ca²⁺-mediated breakdown in red cells indicated that this enzyme is very unlikely ever to be controlled by Ca²⁺ in healthy cells¹⁵. Recent studies of liver and parotid cells have, however, revealed a second type of PI4,5P₂ breakdown that is very rapid, closely coupled to receptor occupation and not Ca²⁺-mediated (refs 8, 16, 17 and unpublished data of L.M. Jones and C.P. Downes). This raises the novel possibility that the initiating reaction in Ca²⁺ mobilization is receptor-controlled breakdown of PI4,5P₂, a characteristic lipid of the plasma membrane, and that the disappearance (that is, apparent 'breakdown') of PI so far studied is really a reflection of the consumption of PI during resynthesis of degraded PI4,5P₂. This does not constitute any fundamental change to the original view that the widely observed 'PI response' is somehow essential to Ca²⁺ mobilization by receptors¹, rather it acknowledges our increasing knowledge of the mechanism of this response: originally we had to discuss the function of stimulated PI and phosphatidate labelling, later we were able to focus on PI 'breakdown', and investigation of receptor-controlled PI4,5P₂ breakdown may now take us one step closer to a final understanding of this long-standing puzzle. The figure shows our current working hypothesis.

In no sense should the information that appears to link breakdown of PI4,5P₂ (or PI) to the mobilization of Ca²⁺ in stimulated cells be taken as conclusive proof of this link. We are still much too ignorant of the details of the processes to draw any such conclusion. Equally, however, I consider that the evidence marshalled as a purported disproof of this hypothesis is far less convincing than Hawthorne³ and Cockcroft² would have us believe. Let us not reject this working hypothesis prematurely. □

1. Michell, R.H. *Biochim. biophys. Acta* **415**, 81 (1975).
2. Cockcroft, S. *Trends pharmac. Sci.* **2**, 340 (1981).
3. Hawthorne, J.N. *Nature* **295**, 281-282 (1982).
4. Putney, J.W. *Life Sci.* **29**, 1183-1194 (1981).
5. Berridge, M.J. *Molec. cell. Endocr.* **24**, 115-140 (1981).
6. Exton, J.H. *Molec. cell. Endocr.* **23**, 233-264 (1981).
7. Irvine, R.F., Dawson, R.M.C. & Freinkel, N. in *Contemporary Metabolism* Vol. 2 (ed. Freinkel, N.) (Plenum, New York, in the press).
8. Michell, R.H., et al. *Phil. Trans. R. Soc.* **B296**, 123-137 (1981).
9. Lapetina, E.G., Billah, M.M. & Cuatrecasas, P. *J. biol. Chem.* **256**, 5037-5040 (1981).
10. Rittenhouse-Simmons, S. *J. biol. Chem.* **256**, 4153-4155 (1981).
11. Farese, R.V., Larson, R.G. & Sabir, M.A. *Biochim. biophys. Acta* **633**, 479-485 (1980).
12. Cockcroft, S., Bennett, J.P. & Gomperts, B.D. *Biochem. J.* **200**, 501-508 (1981).
13. Fisher, S.K., Frey, K.A. & Agranoff, B.W. *Abstr. ISN Satellite Meet.*, Birmingham, 74 (1981).
14. Michell, R.H. in *Phospholipid Metabolism in the Nervous Systems* (eds Horrocks, L.A. & Porcellati, G.) (Raven, New York, in the press).
15. Downes, C.P. & Michell, R.H. *Biochem. J.* **202**, 53 (1982).
16. Kirk, C.J. et al. *Biochem. Soc. Trans.* **9**, 377-379 (1981).
17. Weiss, S.J., McKinney, J.S. & Putney, J.W. *J. biol. Chem.* (submitted).
18. Michell, R.H. *Trends biochem. Sci.* **4**, 128-131 (1979).

X-chromosome inactivation and the control of gene expression

from Martin H. Johnson

WHAT is the relationship between the state of chromatin and the expression of genes? The heterochromatic and inactive X-chromosome appears to offer an appealing model system to tackle this question, since during X-inactivation (or reactivation) the alteration of potential genetic activity, and its inheritance through many cell generations, involves simultaneous changes to a large and coherent group of genes rather than to a series of genes scattered throughout the genome. During the course of a recent meeting in Bangalore*, it became clear that this assumption was too simple. It emerged that the condensation of the X-chromosome, and indeed of heterochromatin elsewhere in the genome, may be a secondary phenomenon, perhaps related primarily to the evolution of sex, and that the selective activity of genes during differentiation might be regulated by other mechanisms altogether.

In the eutherian mammal, the embryo begins life with two active X-chromosomes, but the paternally derived one is later inactivated in trophectodermal and primary endodermal cells (N. Takagi, University of Hokkaido). The cells resemble those cell lineages in adult marsupials in which the paternal X-chromosome is inactive (it is partially active in some lineages), although it is not clear whether this is a result of failure to activate the paternal X-chromosome at fertilization, or of activation followed by preferential inactivation (J. Graves, LaTrobe University). An impressive array of cytogenetic, biochemical and embryological evidence suggests that, in the primary ectoderm (epiblast) of eutherian mammals, random rather than selective inactivation of the X-chromosome occurs between 5½ and 6½ days (Takagi; M. Monk, MRC Mammalian Development Unit, London; M. Lyon, MRC Radiobiology Unit, Harwell).

The simultaneous occurrence of X-inactivation, whether random or paternal and in embryo or embryonal carcinoma cell (Takagi), and the generation of differentiated cells from a stem cell population, lead Monk to propose that both arise from a single genetic change. She suggests that X-inactivation may be just one part of a more general process of genetic determination and that both events might be mediated by the same molecular mechanism. Experiments in which X-inactivation is reversed complicate this simple but attractive idea.

Three types of reversal were described. First, A. McLaren (MRC Mammalian Development Unit, London) presented evidence that the germ cells, although stem cells, had an inactive X-chromosome during their mitotic proliferation in the primitive mouse gonad of 11½ days gestation. Biochemical, genetic and morphological evidence showed clearly that between 12½ and 13½ days both X-chromosomes became active at some point before entry into prophase of the first meiotic division. Second, T. Mohandas (University of California, Los Angeles) exposed mouse-human cell hybrids (*Science* **211**, 393; 1981), in which the human locus for hypoxanthine phosphoribosyltransferase (*hprt*⁺) was on an inactive X-chromosome, to 5-azacytidine, an agent known to reduce DNA methylation (see below), and selected in HAT medium. Clones of hybrid cells, in which human HPRT activity was stably expressed for over six months in culture, were generated and extracted DNA was shown to transform HPRT-negative cells. Graves similarly found that 5-azacytidine treatment of an F₁ cell line derived from a hybrid *Mus musculus* × *M. caroli* (in HPRT-deficient clones of which the *M. caroli* X-chromosome was inactive) resulted in stable re-expression of the *M. caroli* form of HPRT. Loss of HPRT activity in thioguanine-resistant revertants accompanied loss of the *M. caroli* X-chromosome, confirming that the reactivated locus was still on the X-chromosome. In six per cent of reactivating clones, the *M. caroli* form of glucose 6-phosphate dehydrogenase (G6PD) was also expressed, suggesting that reactivation events at the two loci were not independent.

Finally, Graves described experiments in which the *hprt*⁺ locus on the inactive *M. caroli* X-chromosome was reactivated simply by cell hybridization, either with undifferentiated embryonal carcinoma stem cells or with male-derived mouse or Chinese hamster fibroblasts. Again, the loss of HPRT activity in segregants was concordant with loss of the *M. caroli* X-chromosome. The reactivation frequency induced by fusion was much lower than that induced by 5-azacytidine. Recently, de Jonge et al. (*Nature* **295**, 624; 1982) have also reported reactivation of the inactive human *hprt*⁺ gene independent of 5-azacytidine treatment; DNA from human cells carrying an inactive *hprt*⁺ locus was used to transform HPRT-

*The conference on "Condensed Chromatin and the Human X-Chromosome" was organized by H. Sharat Chandra at the Indian Institute of Science, Bangalore, India, 14-16 December 1981.

Martin H. Johnson is in the Department of Anatomy, Downing Street, Cambridge CB2 2DY.

negative mouse fibroblasts and expression of the *hprt*⁺ gene was detected.

In experiments involving reactivation either by 5-azacytidine or by cell hybridization, the inactive X-chromosome remained heterochromatic and allocyclic. Although HPRT activity was being selected for, G6PD and phosphoglucokinase activities were also expressed, but only *occasionally*. A selective reactivation of discrete loci thus seems to have occurred — a process quite distinct from the natural reactivation described by McLaren in which the entire X-chromosome ceases to be heterochromatic and expression of any or all loci can occur. This selective reactivation is reminiscent of the continuing activity of X-linked loci *Xg* and *Sts* on otherwise inactive, allocyclic and heterochromatic X-chromosomes. The X-chromosomes may thus be inactivated at the level of individual genes (which can be switched on or off as can any developmental genes on autosomes) or at a higher level where the chance of individual genes switching on is reduced. The use of 5-azacytidine or cell fusion for reactivation may somehow selectively penetrate the higher-order suppression. That conclusion made it seem less likely that X-inactivation will prove to be a good model for studying the control of gene activity during development.

The capacity of 5-azacytidine to induce reactivation of X-linked loci is consistent with a model for X-inactivation first put forward in 1975. R. Holliday (National Institute for Medical Research, London) and A. Riggs (City of Hope Medical Center, Duarte, US) suggested that methylation of CpG sequences at replication would provide an inheritable change in the DNA that could be associated with suppression of activity and hence could also lead to X-inactivation. Suppression of methylation during one round of replication would generate hemimethylated strands which, on a second round of replication, would leave some cells carrying non-methylated sequences and other cells methylated sequences. Divergence of gene activity in different cell types could thus be achieved. Two rounds of DNA synthesis do indeed seem to be required between each determinative event during early mouse embryogenesis (M. Johnson, University of Cambridge).

Reactivation by 5-azacytidine may occur because it substitutes for cytidine but cannot itself be methylated. However, Jones and Taylor have observed that the extent of demethylation induced by 5-azacytidine is significantly greater than would be expected from substitution alone, and may thus involve a more generalized suppression of methylase activity. If demethylation by 5-azacytidine does occur on such a large scale, why is the whole X-chromosome not reactivated instead of just one or two loci? It seems clear that methylation changes are not in themselves adequate to explain the phenomena

observed — a conclusion supported by a recent report of Wolf and Migeon (*Nature* **295**, 667; 1982). They compared the methylation patterns of DNA from cells with active or with active + inactive X-chromosomes by Southern blot analysis using cloned X-chromosome-specific probes and found no evidence to support any relationship between X-chromosome inactivity and degree or pattern of methylation.

Exciting news came from K. Jones (University of Edinburgh) who described a probe of approximately 200 base pairs, highly repeated, made to the minor density fraction of satellite DNA quantitatively associated with the W chromosomes of the banded krait, a snake in which the female is the heterogametic sex and the W chromosome is therefore equivalent to the Y chromosome of mammals. Remarkably, the Banded Krait Minor density fraction (BKM) probe also hybridizes to the W chromosomes of the females of a number of species of snakes, either at a single discrete area or at a series of sites along the chromosome; to bird W-chromosomes (female heterogametic); to regions tentatively indicated to be associated with the transposed and active mating-type locus (*MAT*) of yeast; and to the mouse (male heterogametic) Y-chromosome in the centromeric region. In addition, the probe bound to scattered sites elsewhere throughout the autosomes of mammals and also to mouse X-chromosomes carrying the '*Sxr* gene'. The gene appears to be involved in primary sex determination as, when present in XX or XO mice, it causes a testis to be constructed. The origin of the *Sxr* locus has long been obscure, but Jones' demonstration that the highly conserved BKM target sequence is translocated from the Y-chromosome to the terminal arm of one X-chromatid during meiosis suggests that either the conserved

sequence, or more probably an associated flanking sequence, is the region coding for the sex-determining function of the Y-chromosome.

Jones suggests that this highly conserved sequence, in addition to its association with sex-determining loci, also defines functional domains in chromatin which are heterochromatic and allocyclic, as for example observed in the probe-positive region of the Y- and W-chromosomes, or is involved in the DNA conformational changes associated with conversion of yeast mating types from repressed to active states. The probe also hybridizes in the region of the proximal-most euchromatin (19A-D) and in the adjacent heterochromatin (19E-20B) of the *Drosophila* X-chromosome, a region proposed earlier in the meeting (D.L. Lindsley, University of California, San Diego) to be the site of sequences responsible for the precocious X-chromosome inactivation in the primary spermatocyte.

Jones proposed that the association of the highly conserved sequence with heterochromatic regions may have arisen in conjunction with the evolution of sex. If the sequence occurred on different primitive X-chromosomes to varying degrees, two classes of X-chromosome might develop, depending on the strength of the inactivating centres. The coincidence of strong centres on the chromosome carrying the putative sex-determining gene(s) could lead to genetic inactivity of the primitive X-chromosome at all times *other* than at the moment of determination of the male sex. Thus, most genes would rapidly be lost, generating a Y-(or W-) chromosome. With weak centres on the primitive X-chromosome lacking the sex-determining gene(s), activity would remain. On this basis, X-inactivation might be viewed simply as a relic of the requirement for a sex-determining chromosome. □

100 years ago

THE ORIGIN OF THE SIGNS OF THE ZODIAC

THE origin of the signs of the zodiac is a question which we have but recently obtained materials for answering. Some of these symbols, certainly, are plain enough: it is not difficult, for instance, to discover the horns of the bull in the symbol of Taurus, or the arrow in that of Sagittarius. But the meaning of others, such as the symbols of Virgo, of Scorpio, or of Capricornus, is not so self evident. These symbols, however, are of comparatively modern invention, and first came into use along with the symbols still employed by astronomers to denote the planets. They cannot be traced further back than the tenth century, and owe their origin to the connection the alchemists believed to exist between the planets and the metals.

But modern though the symbols of the planets and zodiacal signs may be, it is quite otherwise with the signs themselves, and the

majority of the names by which we still call them. Recent research has shown that the general voice of classical antiquity was right in regarding the Chaldeans as the first to map out the path of the sun during the year into separate regions, or constellations. Copies made by Assyrian scribes of older Babylonian works on astronomy have been found in the library of Nineveh, and are now in the British Museum.

It was the primitive population of Babylonia, now known by the name of Accadians, who first made Chaldea famous for its study of astronomy, and it is to them that the Signs of the Zodiac are due. Each sign represented a month of thirty days, and the signs and months were accordingly called by common names.

Some names clearly originated in the peculiarities of the season when the sun was in a special sign of the zodiac. This is certainly the case with Aquarius, and it is probable that fish were particularly abundant under Pisces when the lowlands of Babylonia had been inundated by the rains.

From *Nature* **25**, 525, 6 April 1882.

SPRING BOOKS SUPPLEMENT

Where there's a need for some style

In science-writing it is hard to combine accuracy with fluency. Here Eric Ashby calls for more trouble to be taken over literary style in government reports.

IN St Matthew's Gospel it is written: "Thou shalt love thy neighbour as thyself". A more recent writer puts it another way: "It is desirable that there should be established a meaningful and responsive reciprocal personal relationship within the social group". Here the choice between plain prose and jargon is easy to make. But some messages cannot be coded into words anyone can understand. Every issue of *Nature* carries pages that are meaningless except to a coterie of experts. This is not because the articles are badly written. It would be unreasonable to expect the author of an article on "Sub-GeV antiprotons in galactic cosmic rays" to make his message intelligible to anyone who can read a newspaper. He can assume that the few score of fellow experts for whom he is writing regard the content as almost all that matters in the article. He must avoid inaccuracy and ambiguity; fluency and elegance in style, though welcome and likely to enlarge the circle of readers, are of secondary importance.

The science-writer who chooses a wider readership has a more difficult task. No one is obliged to read what he writes. His motive — to communicate or maybe just to make money — will be defeated if he doesn't take trouble to set out his ideas in lucid, even if he can't manage limpid, prose. He has to adopt a style acceptable to the market place. This exposes him to certain temptations.

The chief temptation is to strive to be lucid at the cost of being accurate. How bracing it is to breathe the fresh air of Macaulay's prose: and how mistaken to suppose that science can be written in that style. It has been unkindly said about Macaulay's style of writing that you can never tell the truth in it. "In satisfying his passion for clarity" wrote the philosopher Brand Blanshard, "he allows himself to omit shades and qualifications that are there in the facts, but would smudge his sharply etched lines if he were to put them into the picture".

The science-writer with a passion for clarity has to take care that his passion does not lead him into infidelity to truth; for the reader who at first takes pleasure in the "sharply etched lines" may in the end feel resentment at having been deceived. That is what befell the American authors of *The Limits to Growth*. They arrest the reader's attention with a vivid scenario, such as the one that pollution, if it gets worse, may almost halve the expectation of life in industrial societies. They take no account of the fact that industrial societies react strongly to any threat to the expectation of life — as a dozen laws to abate hazards to health bear witness — and the arresting scenario is in fact deceitful. It is commonly said that without the "sharply etched lines", the reader would not have taken the trouble to read the book at all. This is the dilemma, and the fact that so few writers can resolve it (and so many do not even try) provokes scientists to frown on a colleague who goes in for popularization; the word is so uncomfortably close to the French *vulgariser*. The rare writer who can successfully combine accuracy with fluency is not frowned upon: he is envied. But for most of us who write about science it is safer to settle for being painstaking about accuracy and be content with a modest measure of fluency.

However there is one category of science-writing where there needs to be a campaign for more fluency, even for elegance, in style: that is in the reports of scientific committees set up by government. The purpose of these reports is to help to shape science-policy. The committees that draft them take immense trouble to make them trustworthy. But — with some outstanding

exceptions — little trouble seems to be taken to make them attractive to read. The consequence is that some reports are read only by those whose duty compels them. And no wonder: listen (it's better to read them aloud) to a couple of examples taken at random from two important reports. Here is a paragraph from the recommendations of the Finneston Report, *Engineering our Future*:

All those involved with manufacturing industry, whether directly or indirectly, should review their activities to ensure that they perceive and present engineering and engineers as matters of vital national concern in their own right.

Does this mean: "The importance of engineers for industry and for the nation needs to be stressed"?

And here is a passage from the recommendations of the report on *Global Resources* to the President of the United States:

Because of the risk to peace, the United States should encourage other countries, and should itself, establish conflict resolution arrangements which can mitigate or help resolve future international water conflicts.

Does this mean: "There is an urgent need for ways to resolve conflicts over competing demands for water"?

There is no need for reports to governments to be written in this stilted and stumbling style. There are models of elegance on both sides of the Atlantic. The report, *A Race against Time*, by the Royal Commission on Electric Power Planning in Ontario, and the report by Mr Justice Berger on the Mackenzie Valley Pipeline (*Northern Frontier, Northern Homeland*), make compelling reading. These are both from Canada, where there is a higher level of literacy in government reports than in the USA. In Britain we have the example of an experiment in fluent writing with, so to speak, a "control" issued in the same year by the same government department. In 1970 the Ministry of Housing and Local Government issued two reports, one from the Technical Committee on the Disposal of Solid Toxic Wastes and one from the Working Party on Sewage Disposal. Neither of these topics is likely to inspire felicity in style, but the contrast between the two reports is striking. That on solid toxic wastes is pedestrian, the format dull, the readership correspondingly small. The report on sewage disposal was a pioneer in the attractive packaging of government reports. The photographs on the cover, the clever diagrams, the homely style, all these say to the reader: "This is not written just for civil servants: it is for you".

Stimulated by pressure groups and the media, the man-in-the-street is becoming more and more concerned about technical decisions which affect his life. If he cannot participate in the decision-making (in public inquiries and the like) he at least wants to understand on what evidence the decisions are made. It is therefore a prime duty laid on those who write the reports of advisory committees and commissions to make them attractive and lucid enough to appeal to a much wider circle of readers than a handful of politicians and civil servants. It would be a public service if reviewers of such reports were to pay attention to style as well as content, praising good writing when they find it and censuring prolixity. Whitehead wrote that it is more important that a proposition be interesting than that it be true. Reports from scientific committees to governments need to be both interesting and true. □

Heavenly insemination

Hugh Montefiore

I SUPPOSE that it is no more strange that an outstanding biologist such as Francis Crick should give his views on cosmology than that a distinguished astronomer such as Fred Hoyle should pronounce on molecular biology. To an outsider, these are both welcome pointers towards an integration of the sciences. So perhaps it was not outrageously odd to give Crick's book to a theologian for review; but he must walk warily on alien territory.

Crick writes ostensibly to answer Enrico Fermi's famous question, "If there are intelligent beings in the galaxy, why aren't they here?", and he assesses the hypothesis known as Directed Panspermia, that is to say, a variant of Arrhenius's nineteenth-century theory, modified in that Crick considers whether life was *deliberately* planted on Earth. His book has nothing in common with Hoyle's views propounded in *Evolution from Space* (Dent, 1981). Crick, for example, does not seriously question whether the period needed for life to emerge by natural process requires a time span longer than the age of the Earth. Nonetheless he seems somewhat sensitive to his wife's criticism that his work may seem more like science fiction than a Nobel Prize winner's critique of a scientific theory. Perhaps for this reason Crick takes the opportunity (as his title suggests) to give his view on life itself, its origin and nature. Written with enviable simplicity and free of abstruse technicalities, it is as a result devoid of any annotation.

The answer that Crick gives to Enrico Fermi is that life on Earth could well have originated elsewhere in the Galaxy, and that there has been time enough for intelligent beings to evolve elsewhere in a suitable environment and to have dispatched prokaryotic and eukaryotic microorganisms by rocket to this planet, whence life here may have developed. Crick, however, admits that this theory of Directed Panspermia, although plausible, suffers from extreme paucity of evidence. He calls it "premature". Crick the human being affirms: "Once the scale and nature of the galaxy is appreciated, it is intolerable not to know whether we are its sole inhabitants"; but Crick the scientist is cautious: "I cannot myself see just how we shall ever

Life Itself: Its Origin and Nature. By Francis Crick. Pp.192. UK ISBN 0-356-07736-5; US ISBN 0-671-25562-2. (Macdonald/Simon and Schuster: 1982.) £7.95, \$13.95.

decide how life originated".

A small measure of support for Directed Panspermia is seen in the almost uniform nature of the genetic code. Some factors, however, Crick does not consider; for example, the origin of some bacteria with properties which seem unlikely to have

some sense, the origin of life appears at the moment to be almost a miracle.

I wonder, however, whether Crick's view will continue to prevail that "it seems almost impossible to give *any* numerical value to the probability of what seems a rather unlikely sequence of events" leading to the emergence of life. In any case the really fundamental problems remain to be solved; how DNA, RNA and enzymes were originally formed, and how the first cell, with its reproductive mechanism, came into existence.

Whether all the dense and sober argumentation of Crick can be validated, I am not in a position to judge. Such matters as available environments for life to emerge elsewhere, the length of time involved, the mode of travel and the duration of the voyage cannot be more than scientific speculation. Crick's scientific writing is lucid, magisterial and suitably cautious ("seems to be", "apparently", "so far as we know"). But in Chapter 15 ("Why Should We Care?") the tone abruptly alters. Crick the Nobel Prize winner has given way to Crick the human being. He reminds me here of those days in Cambridge when he offered a cash prize for the best secular use of a college chapel. When it comes to religious beliefs, gone are the qualifications and dogmatic certainty takes over. "Most modern scientists do not subscribe to *any* of them" (has he carried out a scientific survey?); and

The plain fact is that the myths of yesterday which our forebears

regarded not as myths but as the living truth, have *collapsed*, and while we are uncertain whether we can successfully use *any* of the remaining fragments, they are *too rickety* to stand as an organized interlocking body of beliefs [italics mine].

Francis Crick's personal convictions will doubtless be read with respect; but he would put us in his debt if he would give some explanation of their logical connection with the natural sciences of which he is such a pre-eminent practitioner. Wherein, for example, lies a scientist's "almost boundless optimism concerning his ability to forge a new set of beliefs"? It might seem to some, that the dreadful threat of thermonuclear war overhanging



evolved to meet our earthly conditions. And it is surprising that he does not mention the microfossils which suggest that eukaryotes existed at least as early as prokaryotes, although the initial gap of a billion years means that we do not know how either evolved. As for the supposedly random emergence of life in the prebiotic broth, Crick writes:

It is impossible for us to decide whether the origin of life here was a very rare event or one almost certain to have occurred;

yet elsewhere he writes of it as "an infinitely rare event" and even admits that

an honest man, armed with the knowledge available to us now, could only state that, in

the civilized world, with one-fifth of all scientists at work on defence contracts, should fill us with almost boundless pessimism. Granted, however, "the tremendous success of science, especially in the last hundred years", what is the logical connection between scientific explanation and ultimate meaning, or between scientific knowledge and ultimate belief? Or again, what is the scientific basis for the phrase "outmoded religious beliefs"? For example, is a scientific evaluation of the origin of life compatible with a theistic concept of creation? Readers may wish that Crick had here argued his case rather than merely stated his conclusion. Seen from our anthropic viewpoint, a remarkable series of "coincidences" must have taken place onwards from the initial explosion which brought the cosmos into

being, so as to make possible the eventual emergence of *Homo sapiens* (whether or not ours is the only form of intelligent life in the Universe). These "coincidences" certainly occurred. They may be due to complex sequences of random and meaningless events. But they are also compatible with divine providence. Crick tells us that in considering the origin of life a "gut reaction" is likely to be superficial or misleading. Not all intuitions however deserve to be described in such pejorative terms. Spiritual insights are not to be despised as adequate explanations to such questions which, despite Dr Crick's almost boundless optimism, seem to transcend the fields of scientific investigation. □

Hugh Montefiore is Bishop of Birmingham.

Wholes and parts, meaning and mechanism

Marie Jahoda

The Reenchantment of the World. By Morris Berman. Pp.353. Hbk ISBN 0-8014-1347-0; pbk ISBN 0-8014-9225-4. (Cornell University Press:1982.) Hbk \$43.10, £24.95; pbk \$11.10, £5.95. *The Turning Point: Science, Society And The Rising Culture.* By Fritjof Capra. Pp.464. US ISBN 0-671-24423-X; UK ISBN 0-7045-3054-6. (Simon & Schuster/Wildwood House:1982.) \$17.50, £9.50.

THE revolution in the basic assumptions of physics brought about by Einstein and by the development of quantum theory has repercussions far beyond that discipline. One of the consequences of the new physics — that the mind of the observer is a necessary ingredient of the structure of theory — has not only returned terms like "mind" or "consciousness" to respectable scientific status; more far reaching in the views of a growing number of scientists, philosophers and historians of science it also undermines the mechanistic world view that originated with Bacon, Descartes and Newton.

While physicists gradually absorbed the new ideas in their own field, others spelled out the wider implications of these revolutionary changes for the entire scientific enterprise. Michael Polanyi, for example, argued for the subjective character of all knowledge (*Personal Knowledge*; Routledge, 1958); Michel Foucault in the *Order of Things* (Tavistock, 1966) demonstrated the dominance of epistemes, the communality of tacit basic assumptions in the most diverse sciences at given historical periods — in other words, the domination of world-views in science and their periodic changes; Thomas Kuhn, in *The Structure of Scientific Revolutions* (University of Chicago Press, 1962 and 1977), has made

the notion of shifting paradigms fashionable.

To the extent that there is agreement among these authors and with others who write in the same key, all of this amounts to a clarification, perhaps a redefinition, of the aims of science: no longer is science conceived as cumulatively proceeding to the establishment of ultimate truth about a finite number of laws that govern a finite universe; no longer can science make apodictic statements about ontological essence since the laws of nature are not simply out there but anchored in the minds of men. Rather it is a conceptual enterprise resting on unprovable basic assumptions that have changed in the past and presumably can change again.

Physicists who have made their thoughts about the philosophical implications of quantum theory explicit seem to agree with this conception of science. But the mechanistic view still dominates and the search for discovering the real essence of the world continues. Paradoxically, perhaps, the "hard" sciences are now more open to admitting subjective meaning into their theories, while the "softer" life sciences, particularly aspects of biology, progress on the mechanistic road.

The two books here under review continue the debate with several new twists. Their virtually simultaneous publication, though obviously written independently from each other, must be understood as indicating that the lines of thoughts and arguments they present are gaining ground. Otherwise their truly amazing similarity would be hard to explain. Both have the same scope, beginning with the scientific revolution, proceeding to a careful analysis of the thoughts and contributions of Descartes and Newton, and ending with a sketch of

quite similar new and gentler world-views, that do away with the body-mind dichotomy and the idea of the control of nature in favour of an ecological perspective in which the distinctions between subject and object, value and fact are deliberately blurred. Both authors are deeply impressed by what they see as a convergence with Eastern philosophy (Yin-Yang, Zen Buddhism) of the ideas in quantum physics; both condemn the excesses of reductionism in favour of holism; both write for the general reader (otherwise I would not dare to review them) but in a scholarly fashion, with detailed notes and references; both regard Gregory Bateson's *Mind and Nature* (Dutton, 1979) as the most seminal modern work; both have the same critically respectful attitude to Freud, while preferring the development of his ideas by Carl Jung, Wilhelm Reich, Ronald Laing and Herbert Marcuse. Finally, both are centrally concerned with describing the misery of the modern world with its violence, poverty, consumerism, alienation, inflation, drug dependence, lunatic arms race and ecological destructiveness. In this last respect they are, of course, not alone. In contrast to the Meadows' *Limits of Growth* (Universe Books, 1972), with which they have otherwise much in common, they are cautiously optimistic because they see the new scientific world-view already emerging in North American culture.

Differences between these two books are largely matters of emphasis. Capra, a physicist, is particularly good in describing the development of physics from Newton to the present and is careful to point out which of the modern theories he adheres to are still controversial, for example, Chew's S-matrix theory. While he represents it in a language understandable to the layman and admits that he singled it out because it supports his broader views, it remains impossible for the non-physicist to understand its significance. Berman, a historian of science, writes most interestingly about Newton's psychological conflicts and complexities, and goes altogether more into psychological and political matters. While he, as much as Capra, admires Bateson whose thoughts and sources he describes in greater detail with particular emphasis on the principle of the essential incompleteness of knowledge, he also recognizes that some of Bateson's concepts are double-edged and could be exploited for totalitarian and anti-intellectual purposes; of which he strongly disapproves.

Two major common themes in these books deserve, however, critical comment. The first is the postulated link between the current malaise in the industrialized world and the mechanistic assumptions of Newtonian science; the second and related theme is the reductionism-holism issue.

A yearning to make coherent sense out of one's experience of being alive is probably universal; in Capra and Berman,

both equally alive to the crisis-ridden state of the modern world and the state of modern science, their search for coherent meaning leads them to fuse these two realms into a world-view in which they are inextricably interdependent. Theirs is one interpretation of the past course of history and its possible future trend, but in the light of their own words they are honour-bound to admit that it is a subjective interpretation and that others are possible. One cannot take exception to their yearning for a better world nor to their description of the development of science. But there are good reasons to regard the link between them as not established. Human ingenuity in creating untold misery did not wait for the development of a mechanistic world-view. When Rome burned while Nero fiddled, when under-occupied soldiers in the fourteenth century turned brigands and plundered, raped and burned the French peasantry, devastating its agriculture, very different world-views prevailed. The holistic world-views that have for thousands of years dominated thought in the Far East have not avoided hunger, violence and overpopulation, nor the Cultural Revolution. Furthermore, the assumption that for the period with which these books deal one world-view permeated societies is questionable. People did not and do not regard themselves as machines, notwithstanding the existence of robots. In their search for meaning they have turned neither to Newton nor Niels Bohr, but increasingly in the United States to curious sects, pseudo-religions, creationism, blind imitation of Eastern practices, extrasensory perception, spiritual media and astrology — all of them dangerous in their thoughtlessness as Berman explicitly recognizes, but all of them nearer to the advocated world-view than to mechanistic conceptions.

If the argument is, however, that it is not the general population but those who hold political and economic power who exemplify the Newtonian system and thereby run us into destruction, one could equally well argue that their many misdeeds are the result of sharing essential ingredients of the new view: all too many political and economic pronouncements blur fact and value, object and subject, just

as Capra and Berman advocate.

Thus one of the principal themes in these two books, the link between mechanistic science and the state of the world, stands on shaky grounds.

The other — holism versus reductionism — is treated by both authors somewhat more gingerly; both are too sophisticated to deny the enormous successes of the reductionist approach in science, but both argue that in science and the world at large holism should now replace it.

To the question of reductionism versus holism Douglas Hofstadter has surely provided the final answer in his marvellous *Ant Fugue (Gödel, Escher, Bach; Harvester, 1979)* with the Zen word "mu" which means: unask that question. It all depends on what one wants to know. If one is after the chemical composition of cells, holism won't do; if one wants to know how people cope with the crises of life, no reduction to physiological brain-processes will provide an answer, even though brain-processes are required in coping with life. The step from meaning to mechanism inevitably avoids the question to which an answer was originally sought. The meaningful whole is indeed different from its constituent parts and must be studied on its level, but unless the parts are properly functioning, wholes would collapse. So the parts, whether of human being, of ant colonies, of the environment or of the planetary system must be understood too.

It is true that many scientists still regard the reductionist approach as more "scientific", whatever that may mean. But the development of systems theories is already a powerful antidote to such narrow scientism; Capra and Berman furthermore quote many good minds (and some not so good) who are already grappling from a system's point of view with the central task of all scientific enterprises: to formulate significant questions and design ways to explore their implications. Whether such questions can be raised and tackled by relying on the poetic and mystical world-view of the unity of nature where everything is interdependent with everything else seems to me doubtful. The oceanic feeling of being at one with the Universe is a wonderful experience in moments of ecstasy. For the more pedestrian rational enterprise of science it must be replaced by asking questions of partial systems, small enough not to transcend the powers of the human mind.

Capra in his preface admits that with so large a scope he may have been superficial and simplistic when discussing many diverse fields of study, but he trusts that the whole of the book will be more than the sum of its parts. In contrast, this reader found many of the parts in both books informative and interesting, but the whole a bit muddled. □

Marie Jahoda is Professor Emeritus of Social Psychology and Consultant to the Science Policy Research Unit, University of Sussex.

Genesis of belief

John Maddox

Fundamentalism and American Culture: The Shaping of Twentieth Century Evangelicalism 1870-1925. By George M. Marsden. Pp. 306. ISBN 0-19-502758-2. (Oxford University Press: 1981.) £11.50, \$19.95.

LAST year's trial in Arkansas about equal time for creationism and evolution may have ended satisfactorily, but why was there in the first place a law for Judge Overton to declare unconstitutional? The wish to find out is a good reason for tackling this intricate but elegant piece of scholarship. Disappointment that there is no simple answer is overwhelmed by the richness of Dr Marsden's tale.

The book is an ecclesiastical history of the multicoloured and often warring denominations of American Protestantism in the half-century that ended in 1925 with the Scopes trial in Tennessee. Full-blown Fundamentalism was then only five years old and, Marsden says, never fully recovered from that defeat. Its end is symbolized by the death, on the Sunday after the trial ended, of William Jennings Bryan (counsel for the prosecution), once nearly President of the United States and afterwards, for a time, Woodrow Wilson's Secretary of State. But H. L. Mencken's savage obituary essay could not also serve as an epitaph for grass-roots Fundamentalism — there was too much of it. "Heave an egg out of a Pullman window and you will hit a Fundamentalist almost anywhere in the United States today."

By Marsden's account, Fundamentalism emerged from the reaction by Baptists and Presbyterians against the liberal theology taking hold in their denominations. But its origins lie in the "dispensational premillennialism" of the nineteenth century — the belief that the preordained history of the world specifies a final period (dispensation) of grace preceded by a second coming, itself the end-point of a period of social and moral degradation arranged by the Prince of Darkness. So the literal truth of the Bible, in particular the Revelation of St John the Divine, is the cornerstone of the faith, while saving souls against the second coming (Moody and Sankey in the United States, Spurgeon in Britain) is more important than social reform which might blunt the prophecy.

I had not known that the Fundamentalists own their name to the twelve-volume encyclopaedia of evangelical thought called *The Fundamentals* that was financed by the Californian oil-baron Lyman Stewart and was distributed free to more than 100,000 religious opinion-formers throughout the English-speaking world between 1910 and 1915. By then, the war between the premillennialists and the liberal theologians was well under way. There had been ructions among the Pres-

● Paul and Anne Ehrlich's *Extinction: The Causes and Consequences of the Disappearance of Species*, reviewed in last year's Autumn Books Supplement by Kenneth Mellanby (*Nature* 294, 41; 1981), has just been published by the UK by Gollancz. The book was originally published by Random House in the United States. Price is £9.95, \$15.95.

● A paperback edition of *Biological Energy Resources*, by M. Slessor and C. Lewis (reviewed in *Nature* 283, 316; 1980), has been published by E. & F.N. Spon, price £6.95.

byterians about a proposed liberal revision of the "Confessions of Faith". Bible schools had sprung up, as had evangelists such as Billy Sunday — "I don't know any more about theology than a jack-rabbit knows about ping-pong, but I'm on my way to glory".

Marsden says that the Great War sharpened this conflict. Some premillennialists knew it to be the prelude to the Armageddon that had been prophesied. Bolshevism (1917) was another sign that the Prince of Darkness was at hand. The liberals, on the other hand, feared that such detachment would undermine the war effort. The social upheaval of demobilization, the financial crisis of 1919, the rise of gangsters and even the spread of cigarette smoking confirmed the premillennialists in their gloomy diagnosis. By 1920, both the Baptists and the Presbyterians were at loggerheads among themselves. The Fundamentalist crusade to purge the denominations of soft theology made the Scopes trial inevitable.

What explains the intensity of this dramatic conflict, centred, to begin with, in the north and not the south? Partly it was the nostalgic response to the final passing of the world order, signalled by the War. But Marsden persuasively explains how the tightly-knit immigrant communities of the early decades could have sensed each unfamiliar idea or unexpected event as a threat to the Protestant society they thought they had joined.

Politically, Fundamentalism was conservative (which is not the same as Republican) and fearful of an outside world seemingly rife with Bolshevism and Socialism. On balance, Marsden argues, Fundamentalism was not anti-science as such: indeed, it claimed to be the only scientific interpretation of the facts (the Bible). But evolution, "A string of guesses strung together", was a different kettle of fish. Marsden agrees that in the 1920s Fundamentalism was, for several interesting reasons, largely an American phenomenon. Perceptively, he asks what is to be made of the Fundamentalism now emerging in Islam and (in a footnote) in Ulster.

Marsden's history throws some light on those recent events in Arkansas even though 1981 is well outside his period. It explains why evolution is a unique challenge to Fundamentalism, and why Fundamentalism persists. For you cannot expect such a strong and recent tradition simply to melt away after a few set-backs in the courts. What Marsden does not for me sufficiently explain is why, in a republic in which church and state are constitutionally so separate, so many people have for so long invested so much energy and passion in such fierce and pointless arguments. "The God of your choice", it seems, demands a hard sell. □

John Maddox is Editor of Nature.

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Cartoon from *The King's Business*, a premillennialist journal, May 1925. At this time the Fundamentalists saw Bolshevism, evolutionism and modernism as part of the same basic threat to Bible belief.

Expressing science

Maeve O'Connor

Writing a Scientific Paper and Speaking at Scientific Meetings, 5th Edn. By Vernon Booth. Pp.48. (The Biochemical Society: 1981.) £2.50, \$6.

WRITING [maketh] an exact man, wrote Francis Bacon. How exact does it make scientists, already expected to be the most exact of men and women? And do they learn through writing to become not only more exact thinkers but, in turn, better writers?

Most editors would agree that although some scientists are superb writers and many are adequate, many more are imprecise, obscure, wordy or just plain careless. This is why journals tend to call for accuracy, simplicity and conciseness in their instructions to authors and why some editors devote hours to showing authors how to reach these goals. Editors and their requests for brevity are nevertheless often blamed for squeezing the life out of the so-called "literature". If blame must be assigned, however, we should direct it instead at educational systems that seem hard-pushed to produce literate graduates, never mind literate school-leavers. Yet scientists in particular, who must communicate their thinking successfully to other people if they are to help either society or their own careers, need some training in the art of presentation. Even in this age of cash crises, universities therefore ought to arrange appropriate courses. At present these are rare, but the return on the small investment needed could be enormous.

Until more courses become available, individuals could help themselves by studying one or two books on writing. Among these, Vernon Booth's booklet probably packs in the most good advice per square centimetre. This began life in 1971 as the Koch-Light Laboratories prize-winning essay. Dr Booth has revised and amplified the fourth edition (1977) and added eight pages on speaking at meetings. He does not claim to provide a complete text on how to write but simply to show readers how to avoid the mistakes he has met in scientific papers. He tells us what to do before writing, when to begin ("early") and how to get going ("Begin with the easiest section"). He gives useful hints on handling the various parts of a paper, and he covers all the commonest problems of literary style and punctuation in 12 succinct pages. In the section on speaking at scientific meetings he continues to administer sound advice, though little of this is particularly new. Like the rest, however, this section is pithily put and should be invaluable to busy scientists. □

Maeve O'Connor is Senior Editor at the Ciba Foundation, London, and co-author of Writing Scientific Papers in English (Pitman, 1978).

Science in society, society in science

C.A. Russell

Between Science and Values. By Loren R. Graham. Pp.449. ISBN 0-231-05192-1. (Columbia University Press: 1981.) \$25.90, £14.40.

SCIENCE is about what *is* — what exists “out there” in the given world of nature. Values, however, are about what *ought* to be. They derive, not from ethically neutral scientific data but from the hunches of a Marx, the ambition of a Hitler, the Ten Commandments or even the *Boys’ Own Paper*. At least, that is how the distinction is often made. On that view it is not surprising that the world between science and values has been frequently thought not to exist. Many people still believe this today, though it is a curious fact that they may arrive at such a conclusion from quite opposite directions. One group would deny the concept any validity because science and values are seen as overlapping aspects of one unified whole, each of which can expand to comprehend the other. There is no more meaningful an area “between” them than there is between the North Sea and the English Channel. Another group, however, would conceive of science and values as such different entities that they must be kept in watertight compartments and, since they must not be mixed, nothing could exist that had a little of each.

In this book such groups are labelled respectively “expansionists” and

“restrictionists”. The former hold that science and values did, and do, have much in common and, while in one sense there was no separate world “between” them, they were capable of the most potent interactions that could have consequences for human history of the profoundest importance. That view, rather than the opposite restrictionist one, informs this wide-ranging analysis of scientific controversy in the twentieth century. Graham begins by an examination of relativity and quantum theory, contrasting the metaphysical attitudes of Einstein and Bohr, casting a quizzical eye at the popularizing work of Eddington, and relating to their own cultures the scientific philosophies of the Soviet physicist Fock and the German Heisenberg. In the life sciences, the speculations of Bergson and Monod are followed by the rise of behavioural psychology and ethology. The author has some extremely important things to say about the cases of sociobiology and eugenics and the problems presented in biomedical ethics.

Historians of science will find much to admire in this book, not least its careful documentation, extensive bibliography and authoritative insights into problems that are rather inaccessible (as the history of Russian physics) or else are very recent. But the issues raised are not merely of academic interest; many of them are of urgent practical importance. While the author has no panacea to offer to those facing the dilemmas posed by genetic engineering, biomedical ethics and the claims of sociobiology, he suggests many useful pointers. And one clear conclusion emerges again and again: the inadequacy of restrictionism. Historical analysis repeatedly denies us the luxury of even contemplating the emergence of either science or values without the other. Expansionism is more conformable to historical reality, though prone to abuse and perversion; one need only cite the biologies of Nazi Germany and Stalinist Russia which were “fraudulent interpretations and misuses of science that could have arisen only in their particular political and social milieu” (p. 255).

By concentrating on recent science, and especially those branches concerned with man, Graham has established his case decisively. From now on behavioural psychology, genetics and other related areas of scientific enquiry will be seen by historians as inseparably involved in value-judgements almost from their conception onwards. This is a useful achievement, though not unexpected. It will be good news to those who over the past few years have assiduously sought to debunk “the myth of value-free science”. They may, however, find some of Graham’s other conclusions rather less welcome. Following

Toulmin and others he views science as a spectrum with the value-laden social and human sciences at one end, but with the largely value-free physical sciences at the other. His carefully documented case-studies offer some support for the view recently expressed (by D. M. MacKay) that the notion of *universally* value-laden science is an illogical extrapolation from the social sciences (where it is undoubtedly valid). Although his main purpose is to discover and demonstrate how science has affected values, the reverse process is considerably illuminated by the author’s narrative, and indeed the whole concept of value-laden science is complex enough to justify some analysis. It can mean at least five things.

First, it refers to values involved in choices made by individuals as to research career, subjects and modes of research, publication and so on, and by institutions as to whether to fund research and, if so, in what ways, and whether to exercise any control over it. In the last connection Graham concludes that a core-area of science should be exempt from such control and elsewhere much hard thinking needs to be done before exercising the control deemed necessary. No one would surely doubt that in this almost trivial sense science is always laden with values of individuals and society.

Second, science may be value-laden in terms of its theories. The author lists cases where the cognitive structure of science has been affected by systems as diverse as Puritanism, *Naturphilosophie*, monism, racism and Marxism, their distinctive values affecting the fabric of scientific thought. The different ways in which genetics developed in Germany and the Soviet Union during the 1930s illustrate particularly well the uptake of cultural values in the development of natural knowledge. His example of Copernicanism is perhaps less well-chosen since it may be doubted whether the decentralization of man was seen clearly by Copernicus as a demotion, whether Calvin ever did oppose the new system (p. 303), or even whether the “mountain of evidence” produced by Galileo was so regarded at the time (p. 362). But the essential point remains that Copernicus made a choice rather than a discovery, and that therefore values must have been present at the beginning. It is not going too far to suggest that most scientific theories are at least partly constrained by social values. That, however, is not to say that there is nothing to differentiate them in kind from (say) theories of politics or literature. The difference lies in the appeal to empirical data, and it is in this third area that the assaults on the concept of value-free science are least successful. Given that “a real physical and biological world exists” (p. 25), empirical data and regularities will also exist independently of the ideology of the observer. Nowhere is this more true than in the physical sciences. Perhaps that is why Eddington could be so

The Strain and the Bond

*She was a melancoli clone,
A prey to moods and shyness,
Who stayed at home, and grew alone —
A virtuous f-minus.*

*She loved another from afar:
She idolized his gender.
She little knew of H-fr
(And less of Waring blender).*

*Responding to her coliform,
And tempted by her colihood,
The H-fr took her by storm
(As willful f-rs would).*

*Beguiled by airy persiflage
(A noxious kind of coltergeist),
She caught a bout of coliphage;
Succumbed to it; and lysed.*

*So, maidens, stay alone — and well —
Remaining blithe and bonny,
While other cultures go to Hell
In coli matrimony.*

This poem is taken from Ralph Lewin’s *The Biology of Algae, And Other Verses*, recently reissued by University Press of America at \$6.75. Another taste of Professor Lewin’s verse appears on p.506.

sure of his restrictionist position, with his limitation of science to just "pointer readings".

The application of science to technology marks a fourth opportunity to seek for values, and they are not hard to find. Decisions as to whether and how to apply science inevitably involve value-judgements, and few would quarrel with that assertion. A point of interest that does emerge from Graham's discussion of medical ethics is that strictly scientific knowledge plays a relatively small part in creating and solving the various dilemmas. The case of amniocentesis is instructive. Medical technology, much more than scientific research, has determined the dimensions of the problem, and the social climate is even more important.

The last kind of way in which science could be said to be value-laden is in its attempted extension and application to social problems as such. In this case it is much more true that science influences values than that it reflects them in its structure and practice. And here, in example after example, Graham demonstrates how readily this happens. He traces the story from Einstein (who, though a prolific writer on social matters, denied that science was directly related to values) to the sociobiologists' conclusion that all values are under genetic control — "the ultimate expansionism". *En route* he has much to tell of the ways in which science was used to justify ethical norms (as with Social Darwinism and capitalism) or even to explain them. He reminds us that several key figures, as Bohr, changed their views with time; he observes that early eugenics, far from being a tool of the Right was regarded in Weimar Germany as a Leftist deviation; and he cautions us against always supposing that "science is forcing changes in our values" (p. 267).

The author makes no attempt to conceal his own values. He emerges as a liberal and humane person, determined to see the best in ideologies with which he disagrees. Naturally some of his conclusions are highly contentious. Does it really follow, for instance, that dialectical materialism, whatever its merits, "has been damaged — probably beyond the point of salvage — by the fact that it is the political doctrine of an oppressive, nondemocratic state" (p. 349)? And is restrictionism always anti-scientific? What about Bacon and his two books of *Nature and Scripture*? And who is he that shall decide the issues of human genetics "without a priori commitments" (p. 255)? Nevertheless we can be grateful for an original, provocative and at times brilliant demonstration of the relevance of historical enquiry to critical issues of our day. □

C. A. Russell is Professor of History of Science and Technology at the Open University. His most recent book, Science and Social Change, will be published by Macmillan, London, later this year.

The second use of nuclear weapons

Laurence Martin

Nuclear Illusion and Reality. By Solly Zuckerman. Pp.154. UK hbk ISBN 0-00-216555-4; US 0-670-51822-0; pbk ISBN 0-00-216554-6. (Collins/Viking: 1982.) Hbk £7.50, \$10.95; pbk £4.95.

THE illusion which Lord Zuckerman mentions in his title, and which his book is chiefly devoted to exposing, is the idea that nuclear weapons could ever be used as a rational means of defence. A nuclear war would, he asserts, almost certainly become total, mutually annihilating, even if it began in a small way, which is itself improbable:

nuclear weapons may well be classified as strategic, theatre, and tactical, but these terms are meaningless if the use of one of them may mean the use of any.

In effect, then, nuclear weapons only serve to deter the use of nuclear weapons and Lord Zuckerman, while believing in the necessity for such deterrence, advocates what is usually called a "no first use" policy. Unilateral nuclear disarmament would expose the West to nuclear blackmail but, for all practical purposes, defence against aggression short of nuclear attacks must be provided by conventional forces. It is to those, argues Lord Zuckerman, in common with many voices now raised in the Trident debate, that our military energies should be vigorously directed. On the British nuclear force, he is a little ambivalent. He scoffs at the idea that British disarmament would serve as a useful example to others and, while regretting the resources diverted from conventional forces, he concedes the British force is an adequate deterrent to nuclear attacks on the United Kingdom. His most original reason for favouring the maintenance of the British and French forces is, however, that as small but effective deterrents they serve to demonstrate the superfluity of the huge forces maintained by the United States and Soviet Union.

There is clearly a good deal of plausibility in Lord Zuckerman's arguments and they are advanced with admirable lucidity. He pokes some shrewd holes in much that passes for strategic wisdom: pointing out for instance that those who cite the small CEP (Circular Error Probable) of missiles usually fail to consider the 50 per cent of warheads that will fall elsewhere, some with really gross errors; that tactical nuclear weapons have been widely distributed without any convincing doctrine for their use having been concocted even to the present day; that small nuclear warheads are not even very effective for many military purposes and that efforts to use them may consequently entail using large and devastating numbers. Nor can one fault his

lengthy explanation of what the effects of large scale nuclear war would be, though we are perhaps not lacking in earlier expositions of this theme.

For those already immersed to some degree in the nuclear debate, however, there are two threads in Lord Zuckerman's book that may be received rather cautiously. One of particular interest to readers of this journal is Lord Zuckerman's repeated attacks on the role — the primary role it often appears — of science and scientists in promoting the "arms race". There is, of course, room for debate as to whether the international military competition that undeniably exists is so unlimited, and so useless, as the metaphor of race suggests. This debate is important, for the answer should influence our readiness to seize on such remedies as are proposed.

So far as the specific role of science is concerned in stimulating the competition, whatever its intensity, there are actually two kinds of scientists in Lord Zuckerman's world. The majority appear to be rather lower grade people, usually called "technicians", who work in military research establishments and foist innovations on the military and the politicians. There is also, however, a small elite of scientific advisers — in Britain including Lord Zuckerman himself, in the United States his friends Jerome Wiesner, Herbert York, Franklin Long and so on — who see the uselessness of it all but whose advice never seems to have been taken when anything wrong was done. One can only admire the resolution with which they nevertheless clung on to the job.

The lesser kind of scientist is apparently the fundamental cause of our military predicament: "At base, the momentum of the arms race is undoubtedly fuelled by the technicians in governmental laboratories and in the industries which produce the armaments". There have been several recent studies of this problem, chiefly in the United States, and it undoubtedly deserves even more thought. Neither militarist nor keen disarmer should tolerate military developments that serve no useful strategic purpose. Doubtless some weapons do get made because they are "there" and sometimes, perhaps, a weapon emerges not as the result of a military specification but because a diversified technological process throws up the necessary components: possibly the cruise missile is such a case, coalescing from disparate developments in guidance, propulsion, image suppression and so on.

But this does not mean that the cruise missile is a useless or necessarily deplorable weapon or that its production is the result of spontaneous generation rather than rational strategic decision. Lord Zuckerman seems to regard the MIRV as such a

New Bioscience Titles from Pitman

Attention and Schizophrenia

Robert D Oades

What are attention and schizophrenia? How can they be studied? This book reviews evidence as to which brain areas and systems may be involved in these phenomena. The objective is to make pharmacologists, psychiatrists, neurobiologists and behaviourists more aware of each others' problems, and to stimulate an interdisciplinary approach.

February/344 pp/Cased
ISBN 0 273 08490 9/£22.00

Molecular Approaches to Evolution

Jacques Ninio

Evolutionary theory is considered in the light of modern advances in molecular biology. The fundamental aspects of molecular processes, used by the cell in order to evolve, are explained.

February/140 pp/Paper
ISBN 0 273 08521 2/£5.95

Neuropharmacology of Insects

Ciba Foundation Symposium 88

This volume brings together 24 specialists in insect neurobiology, neuropharmacology, genetics and physiology. They present a thorough review of the current state of insect neuropharmacology, the position it occupies in pharmacology as a whole, and the directions it might profitably take in future.

April/340 pp/Cased
ISBN 0 272 79652 2/£25.00

Competition and Coexistence of Species

A J Pontin

This much needed book reviews a developing subject for undergraduate teaching, in which no suitable monograph has previously existed and for which ecology textbooks are too brief and incomplete in coverage.

January/110 pp/Cased
ISBN 0 273 08489 5/£8.95

Pitman Books Ltd
128 Long Acre
London WC2E 9AN

Pitman

Circle No.26 on Reader Enquiry Card.

technology- rather than strategy-led growth. Yet we know that the production of MIRVs was the subject of prolonged, agonized and explicit debate. It is surely going too far to absolve the political leadership of responsibility and assign it all back to the "technicians". It is interesting to note that in its latest *Yearbook*, the Stockholm International Peace Research Institute flatly repudiates the "driven by technology" thesis and affirms that the political system gets the weapons for which it calls.

Lord Zuckerman does not make very clear what he believes the main danger of the technology-driven arms race to be. Does sheer quantity of weapons make war more likely? There is no definite statement of this. It is not, apparently, that lesser quantities would do much to limit damage, for Lord Zuckerman is adamant that much smaller nuclear forces than the present ones would serve to wreak disaster. But dangerous Lord Zuckerman certainly believes the present military confrontation to be, and he despairs of survival if the East-West competition is not moderated by arms control before the end of the century. Obviously weapons secured for no logical strategic reason would be a waste of money. Reading more between than on the lines, however, it would appear that perhaps the main danger Lord Zuckerman perceives is that an unjustified reliance on nuclear elements in NATO strategy might lead to the use of nuclear weapons despite the fact that it would be irrational to do so.

Not since the first contradictions of the idea of "massive retaliation" has it been possible to deny the reality of this danger, paralleled as it is, in Kennedy's famous phrase "holocaust or humiliation", by the danger of political defeat if the threat to use nuclear weapons proves hollow. Lord Zuckerman's cry for us to look to our conventional forces, and to recognize that nuclear deterrence requires an adequate underpinning of lesser forces, is amply justified.

It is here, however, that an all-too-brief reference to a second debatable theme in the book should be made. Lord Zuckerman is an advocate of "minimal nuclear deterrence" combined with conventional defence, and almost totally rejects any idea that nuclear weapons might be used in limited ways for "war-fighting" or to introduce an "intra-war" level of deterrence by nuclear means. His one concession is to admit that

it is just possible that one side might not reply in kind, say to a nuclear 'shot across the bows' or to the attempted 'discriminate' use of nuclear weapons, or even to a first strategic strike.

But he goes on to say that in practice retaliation would be unrestrained for psychological reasons and for lack of technically adequate command and control. When citing with disapproval Dr Harold Brown's statement as Secretary of Defense — as an earlier "scientific adviser" Brown gets good marks — that a

flexible strategy should "leave open the possibility of ending an exchange before the worst escalation and damage had occurred, even if avoiding escalation to mutual destruction is not likely", Lord Zuckerman says only "the final ten words of this statement are those that matter".

Whether even "not likely", if accurate — and certainly it is hard to be optimistic about such a juncture — justifies making no provision at all for "controlled" nuclear action in such a contingency is, and is likely to remain, a vexed and probably the most important strategic question of our advancing nuclear age. It is good to have Lord Zuckerman's firm view to ponder. The logic of his position leads to his "no first use" policy. But we should not accept this without first dwelling on two or three difficult questions. What is the effect of a "no first use" policy in a coalition where the main and, if many had their way, the only nuclear forces belong to a distant

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Lord Zuckerman — a view to ponder

ally? If nuclear attack on Western Europe may not be regarded as a "first use" on the United States, should the United States have an identifiable separate level of "retaliation" to deter in such cases? And if a future war in Europe is given a virtually assured "conventional" character, by a "no first use" policy, would the calculation facing a would-be Soviet aggressor — "Can I win a conventional war in Europe" — be dangerously different from the one facing him today — "Can I afford to start a war in which nuclear weapons will probably be used"?

In practice, the balance of deterrence in Europe is probably stable over a wide range of strategic postures, because the stakes are so high and the all-pervasive aura of nuclear fear is so daunting, whatever the explicit strategic doctrines. But in the long run, underlying assumptions about how nuclear deterrence is related to overall deterrence must have profound effects on how states prepare in peacetime and might behave in war. For those who think such questions important, Lord Zuckerman's book offers a provocative and forthright essay well worth the reading. □

Laurence Martin, Vice-Chancellor of the University of Newcastle upon Tyne, was Professor of War Studies at the University of London from 1968 to 1977. He delivered the 1981 BBC Reith Lecture, "Armed Force in the Modern World".

Quantum theory: certainty in uncertainty

P. W. Atkins

The Cosmic Code: Quantum Physics as the Language of Nature. By Heinz R. Pagels. Pp.370. ISBN 0-671-24802-2. (Simon & Schuster:1982.) \$17.50.

THIS is a book aimed squarely at the general reader with the intention of conveying not only the excitement of modern science but also the revolution in our comprehension of the world that accompanied the introduction of quantum theory. There are three parts: an account of quantum theory; a description of modern particle physics; and a brief personal reflection on the nature of physical laws. The material is presented non-mathematically and will be accessible to anyone who is interested in what the author rightly regards as the nonpareil of this century's social and cultural events. The text itself is characterized by the author's almost boyish eagerness to communicate his enthusiasm, and is sometimes so anxious to please that it trips over its own naivety ("Most physicists enjoy the outdoors") to the point, in one place, of ascribing the well-known allegory of Paley's watch to a speaker at a recent conference.

The most important section is the first, which occupies about half the book. It is woven around the biography of Einstein and his transition from being a revolutionary innovator to one who lost "his hot-line to the Old One" and refused to accept that determinism was dead. The intention of the section, apart from emphasizing a moral, is to present an exposition of quantum weirdness; the loss of objectivity, its replacement by an observer-created reality and the rejection of determinism — "the quantum theory is a theory of an instrumentally detected material reality" (p.99). Underlying (if not stemming from) this intention is an opposition to material reductionism (a programme that "cannot be carried out", p.135). The basis of this view is an acceptance of the Copenhagen interpretation of quantum mechanics with its emphasis on the meaninglessness of a concept until its mode of measurement has been defined. Pagels adopts the party line on this issue, and as such does a very good job. His argument leads, convincingly, to the ultimate quantum weirdness that "human intention influences the structure of the world" (p. 95), and the view that "quantum reality" is statistical.

Convincingly, that is, until one stops to think. In my view, for what it is worth, determinism entered physics with the uncertainty principle, and the principal confusion pervading so much interpretative commentary on quantum mechanics (and finding such lucid expression in this book) is the failure to distinguish between measurement and

specification. I should explain what I mean. With quantum theory we first encountered a constraint on our description of the world. Classical mechanics attempted unknowingly to be overcomplete in its specification; quantum mechanics is the first theory we have (and possibly the last we shall need) that respects nature by refraining from imposing a description too strong for it to support. Quantum mechanics is fully deterministic in the sense that the evolution of a state under the influence of a hamiltonian is perfectly well-defined; it ceases being deterministic only when we insist upon reverting to an overcomplete discussion by asking questions that classical mechanics has conditioned us to believe are answerable or when, after we have prepared a system in an eigenstate of one observable, we seek to predict the outcome of another observation that classical mechanics has conditioned us to expect but which quantum mechanics proscribes. No wonder our predictions then squirt out in all directions! Quantum "randomness" is nothing more than this reaction to false expectation. Likewise, the strict Copenhagen requirement of "meaningless until measured" is a confusion between a mode of interpretation and a manner of exhibiting self-consistency (or at least of constructing thought experiments to show that there is no inconsistency). This viewpoint is consistent with Einstein's whose views, I suspect, have been overborne by the appeal of an interpretation that hangs on to classical modes of thought. This emphasizes yet again the delicacy of the balance in the dictum (p.67) that "physicists are conservative revolutionaries . . . pseudoscientists lack that commitment to existing principles".

My position then is strict determinism (or better, neodeterminism, because it must not be confused with classical determinism) showing itself as indeterminism only when false, culturally conditioned, classical questions are asked. Hence, in my view, Pagels has focused on a wholly misleading and fallacious interpretation which cannot fail to imbue his readers with a false impression of the nature of "quantum reality". I have to admit, though, that he has the forces of neoconvention ranged on his side (which does not mean he is right), and readers of this engaging book will leave it with a clear grasp of a conventional and widely accepted interpretation.

I have concentrated on a single aspect of the book, but one that lies at its core. There are many other remarks with which the careful reader will probably disagree. For instance, I could criticize at equal length the remarks to the effect that differences between macroscopic and microscopic

phenomena are qualitative and not merely quantitative (p.128), that historical events are not reducible to individual acts on the part of human beings (p.131), and that quantum reality is statistical and therefore outside mathematics (p.337). There is also a peculiar blindness to the technical meaning of a "selfish" gene but not to its analogue "charm". There is also much to admire, especially the author's manifest deep concern to communicate. □

P. W. Atkins is a Fellow of Lincoln College, Oxford, and author of The Creation (W. H. Freeman, 1981).

Getting JET off the ground

John B. Adams

A European Experiment: The Launching of the JET Project. By Denis Willson. Pp.181. Hbk ISBN 0-85274-543-5; pbk ISBN 0-85274-549-4. (Adam Hilger/Heyden: 1981.) Hbk £10.50, \$23; pbk £6.95, \$15.50.

EVERYBODY who is concerned in Western European affairs should be interested in this account of "a European experiment". Opponents of the European Community system will find in it ample confirmation of their criticisms; supporters will see in it a triumph of the European spirit over national self-interest. For those who are convinced that international action is now the only way of carrying out major scientific enterprises in Europe, it will come as a sobering reminder of how very difficult it is to reach agreement at the European level.

The subject of the book is the launching of the Joint European Torus (JET) project, the latest and largest experiment in the world-wide effort which started back in the 1950s to exploit nuclear fusion reactions as a useful energy source. To make this dramatic story (or lamentable comedy as it was once described) intelligible to the layperson, the author has first to explain something about plasma physics, the general history of nuclear fusion research and the complexities of the European Community system, all of which he does with commendable clarity. Having provided the reader with the essential background information, he then describes the specific problems presented by JET.

The principal difficulty that arose in the launching process and the reason for the two-year delay in the approval of the project was the problem of where to locate it. After the failure of the Community to agree to a common nuclear fission reactor programme and the consequent disarray this caused to the Euratom laboratory at ISPRA, which was set up to carry out that programme, the Commission devised a



JET Joint Undertaking

JET takes form — the eight limbs are part of the transformer core of the machine, currently under construction in the Torus Hall at Culham.

different method for pursuing nuclear fusion research in Europe. This was to set up a coordinated research programme using the existing national fusion laboratories rather than to try to bring all the work together in a single European laboratory. This method proved acceptable to the national laboratories and to the Governments concerned, and it worked very well until the JET experiment came along. Because of its large size and the need to engage most of the national laboratories in its design and construction, it did not easily fit into the established pattern. A new system was adopted, called a Joint Undertaking, which made JET a common European enterprise. But there was no European fusion laboratory in which to build it.

The Commission first suggested that JET should be built at ISPRA where there were resources then under-utilized. This proposal was opposed by many of the fusion scientists who feared that the project would be seriously handicapped by the absence of fusion expertise at ISPRA and by its previous history. It was also opposed by several member states, who saw JET as a desirable acquisition and who proposed alternative sites for it on their own territories. The scene was thus set for a first-class battle for the JET site in which all of the many levels of the Community became involved, and it is this battle which forms the main drama of the book.

Denis Willson describes these events, blow by blow, with great objectivity and with refreshing frankness which makes his book a very valuable account of the trials and tribulations of those who have the need — and the courage — to launch international projects in Europe. It may even be used to avoid similar problems arising again in the future but, as the author points out, the Governments in Europe still have no agreed way of resolving such problems and it is difficult to see how they will acquire one as long as they insist on unanimity (which, incidentally, they had finally to forego to reach a decision on JET). But, despite all the difficulties, JET was finally approved and is now under construction next to the British nuclear fusion laboratory at Culham where it appears to be going very well.

Similar dilemmas faced by CERN, the European Organization for Nuclear Research, are occasionally quoted for comparison but it is not mentioned that it took CERN much more than two years to reach a decision on where to build its SPS machine. This problem was finally resolved

when the scientists involved agreed to reshape that project and themselves proposed the only site where it could be built — something which the scientists in the national fusion laboratories were apparently unable to do for JET.

This book is a well-written and thought-provoking account of a European experiment in all senses by someone who played a key role in the launching of JET and who had access to all of the documentation. It can and should be read and digested by all serious-minded citizens of Europe — politicians, civil servants, scientists and laypersons — and it will surely be read and noted by similar people in other regions of the world. □

Sir John Adams, currently at CERN, was Director of the Culham Laboratory from 1960 to 1967.

Space for science?

John Noble Wilford

Beyond the Atmosphere: Early Years of Space Science. By Homer E. Newell. Pp.497. ISBN 0-80-607146. (NASA Scientific and Technical Information Branch: 1981.) Hbk \$12.50; pbk \$9.50.

FROM the start, in 1958, the National Aeronautics and Space Administration and the scientific community have had an uneasy and ambivalent relationship. It was probably unavoidable. For all their shared interests in exploration, they approached the space age with different priorities and expectations. Scientists saw in NASA a new source of research funding and of opportunities to study Earth, the Solar System and the Universe. While NASA was not ungenerous in its early dealings with scientists, its administrators marched to a different drummer. Given their political mandate, which was to establish the United States as the pre-eminent space power, the NASA managers accorded highest priority to projects aimed at "catching up" with and eventually "beating" the Soviet Union. Scientists had no choice but to accept a secondary role in the Apollo Project, for example, and to live with a relatively small share of the space budgets, now as then. Tensions between them and NASA have been an inevitable consequence.

The potential for conflict has, if anything, risen since the Apollo days. Even at the height of the Apollo mobilization, not a year went by without the initiation of a new space-science project. A small share of ample budgets paid for a lot of science. Post-Apollo budget reductions, however, coupled with inflation and the expense of the space shuttle, have now squeezed space science to the point of near-paralysis. New projects are rare, and existing ones are forever being threatened with deferral or cancellation, a fate that recently befell the Venus Orbiting Imaging Radar Mission and the American half of the International Solar Polar Mission. Only desperate lobbying by scientists, led by Carl Sagan, saved the Galileo Project from extinction this year.

In this context of uncertainty verging on despair, scientists will read Homer E. Newell's *Beyond the Atmosphere* with a sense of nostalgia for the early days of space science. Dr Newell, who retired in 1973 as NASA's associate administrator, bore much of the responsibility for mediating NASA's relationship with scientists so that a vigorous space-science programme was possible. He is a scientist himself who brought to his job considerable experience in managing Government science undertakings. His book is another in a series of histories prepared under NASA auspices, which means that it is, in a sense, "official history", though somewhat more candid than most works of this genre.

Much of the book reviews the origins of space science as a part of NASA's operations and the achievements of NASA-financed research from 1958 to the mid-1970s. Space science, the way Dr Newell uses the term, is the multi-disciplinary pursuit of a knowledge made possible or significantly aided by rockets and spacecraft. The breadth of the field as it evolved is impressive: geodesy, meteorology, atmospheric and ionospheric physics, magnetospheric research ("a genuine product of the space age"), lunar and planetary science, solar studies, galactic astronomy, relativity and cosmology as well as the life sciences and exobiology.

Dr Newell acknowledges that the expanding perspective afforded by space-flight has yet to produce any scientific revolution. But he writes with understandable pride of the "continuing harvest" of knowledge from space science, notably in the earth and planetary fields:

No longer was the geophysicist confined to a study of only one body of the solar system. No longer was the study of the planets solely a venture of the astronomers. The dearth of new data that had led planetary studies into the doldrums and even disrepute . . . gave way to a sudden flood of new information that re-awakened the astronomer's interest.

Other exciting developments included the effect on astronomy of space observations in the hitherto hidden wavelengths and the

discovery of Earth's magnetosphere, not suspected beforehand.

For those scientists who have worked in NASA projects, or expect to, Dr Newell's concluding chapters will be of greatest interest. They deal with "the real world of budgets and finances". They seem at times to be a plea for more understanding by scientists of how competitive are the claims for NASA's favour.

Dr Newell is perhaps still too much the diplomat to render a critical historical treatment of the struggle between scientists and engineers within NASA and between unmanned and manned programmes. Such struggles persist to this day. The author, taking note of bitter feelings among scientists towards the manned programme, does remind them that even if they had succeeded in reducing or cancelling a project such as Apollo, there was little likelihood that much, if any, of the money saved would have been diverted to them. To the contrary, he observes, science generally benefited from Apollo by riding on its coat-tails for such lunar science endeavours as Surveyor and Lunar Orbiter. Moreover, he contends that scientists should have been grateful that they had a civilian agency like NASA rather than having to depend on military support, as had been the case in the pre-Sputnik days.

As for the future, Dr Newell proffers cautious optimism about the place of space science in NASA's operations, or at least he puts the best light on post-Apollo developments. Out of the "searching scrutiny" in the early 1970s, he writes, "emerged an acceptance of a continuing role for the agency in which science, applications, and exploration would all play a part". Moreover, the author says, NASA's position became "intrinsically stronger" because it is now "freed at last from an uneasy dependence on a passing sense of urgency over the nation's technological strength relative that of the USSR".

A failure of the book is the absence of any critical evaluation of the structure for the support of space science as it has evolved. Would space science be served better if it could count on several major sources of funds, not just NASA?

Dr Newell does advise that scientists must scale down some of their expectations if they are to win continuing support from NASA. These words are only now being heeded. Resistance by recent Administrations to large, new projects — the expensive planetary missions in particular — has prompted scientists working with NASA to shift their thinking to more modest goals.

Accordingly, the Space Science Exploration Committee, formed by NASA last year and chaired by Noel Hinners, Director of the National Air and Space Museum, is devising a new strategy. Missions with limited, specialized scientific objectives, employing lower-cost, standardized spacecraft, would be emphasized. Major undertakings, on the scale of another Viking or Space Telescope, would have little or no

place in planning for the foreseeable future.

It remains to be seen how acceptable this modest strategy will be to NASA's managers or the Reagan Administration. A review of the American space programme, being prepared now by the White House Office of Science and Technology Policy, will presumably have more influence on the future course of NASA than any group of scientists. If the Administration's pronouncements are any indication, the review will deal almost exclusively with the space shuttle and policy regarding NASA's jurisdiction *vis-à-vis* the Department of Defense. Meanwhile, the engineers in NASA are drawing up plans for another big project, the manned space station, that would probably absorb any funds released as a result of the shuttle's completion.

This can only give scientists new reason to be uneasy. The entire American space programme is in a state of flux, with the ascendancy of the Pentagon in space affairs and the re-ordering of NASA's goals and responsibilities. It is in times like these that space scientists may find themselves wishing Homer Newell was still in a position to attend to their interests within the chambers of NASA. □

John Noble Wilford is a science correspondent of The New York Times who covers NASA. His most recent book is The Mapmakers (Junction Books/Knopf, 1982).

Heroes of AI

Richard Gregory

Science Observed: Essays Out of My Mind. By Jeremy Bernstein. Pp.376. ISBN 0-465-07340-9. (Basic Books: 1982.) \$16.95.

THIS is a personal account, in 17 essays, of recent highlights of physics and of artificial intelligence. The pieces are written in lively style by Jeremy Bernstein, a professor of physics and author of several well-known books, including *A Comprehensible World* (1961), *The Analytical Engine* (1964), *Einstein* (1973), *Experiencing Science* (1978) and *Hans Bethe: Prophet of Energy* (1980). He also writes on science for the *New Yorker*.

Like ancient sagas, these essays are based on heroes, on men of science confronting the unknown and sometimes each other. Chief among them are Marvin Minsky, Einstein, Robert Oppenheimer, Schrödinger and Harold Furth. Others lie in the shadows, seen with fleeting ambiguities: Ernst Mach — father and disowner of relativism. Charles Babbage — grandfather of artificial intelligence, who set out to weave algebraic patterns with his

analytical engine, just as the Jacquard loom weaves flowers and leaves. Babbage spoke most clearly through the voice of Ada Augusta, Countess of Lovelace, rejected daughter of the poet Byron. And Alan Turing — father of artificial intelligence, who dared to oppose the wit of men against machines, to see if under their disguises men and machine are the same. (Or is oil thicker than blood?)

The first principal hero is the guiding genius of artificial intelligence at MIT, Marvin Minsky. Indeed he is, as I know by acquaintance and not only description, all that Jeremy Bernstein claims. With his equally brilliant colleague John McCarthy, now at Stanford, he started the AI enterprise in the United States in the late 1950s, and they invented its name. "Artificial intelligence" is philosophically better than the rival "Machine intelligence", for if brains are intelligent machines the colours may get confused.

Minsky is a compelling writer as well as a delightful companion and true inspirer of students. Disarmingly, he dislikes the activity of programming, preferring to think and discuss and write in words. Of intelligence, Minsky wrote in 1961 (quoted on p.54): "To me 'intelligence' seems to denote little more than the complex of performances which we happen to respect, but not understand". And, "Programmers, too, know that there is never any 'heart' in a program. There are high level routines in each program, but all they do is dictate that if such and such, then transfer to such and such a subroutine". And, "When we look at the low-level routines, which actually 'do the work', we find senseless loops and sequences of trivial operations, merely carrying out the dictates of their superiors".

There is much of interest here on the history of computing, with figures of the sizes and costs and weights and energy consumptions, as well as the speed and power of ancient and modern computers through their generations. It is indeed astonishing that ENIAC, which in the late 1940s was the most complicated electronic device ever built, was a hundred feet long, ten feet high and three deep. It consumed 140 kilowatts, to handle twenty ten-digit "words" in its memory. On this scale, one could not get a present-day home computer into the house, or even the street.

Jeremy Bernstein is a scientist writing for scientists. He assumes that one knows about Mach's principle and Newton's bucket; and he also comments economically and without gush on such matters as the idiocy of using troops to test the effects of atom bomb blast and how well they would stand up psychologically to nuclear war. These are extremely readable sagas of our age; and they are well worth reading. □

Richard Gregory is Professor of Neuropsychology and Director of the Brain and Perception Laboratory at the University of Bristol.

What skulduggery?

Steve Blinkhorn

The Mismeasure of Man. By Stephen J. Gould. Pp.352. ISBN 0-393-01489-4. (W.W. Norton: 1981.) \$14.95, £9.95.

WITH a glittering prose style and as honestly held a set of prejudices as you could hope to meet in a day's crusading, S.J. Gould presents his attempt at identifying the fatal flaw in the theory and measurement of intelligence. Of course everyone knows there must be a fatal flaw, but so far reports of its discovery have been consistently premature.

The theme of this particular book is that since science is embedded in society, one must expect to find the prejudices of the age presented by scientists as fact. Most authors, given such a theme, would be content to document and catalogue instances in support of the proposition. Gould, however, goes one better by writing a book which exemplifies its own thesis.

It is a masterpiece of propaganda, researched in the service of a point of view rather than written from a fund of knowledge. For the best propaganda requires not the suppression or distortion of facts but their careful selection, emphasis and juxtaposition. So, in a work which declares its concern to be with the notion of intelligence as a single measurable "thing" in the head, we find that two-thirds of the argument is given over to a careful reworking of early attempts to establish craniometric and anthropometric criteria of intelligence, and an admirably disturbing account of the Gadarene rush to press IQ tests into the service of social engineering in the USA in the first half of this century. As Gould rightly emphasizes, many of the uses to which tests were put made mockery of their original purpose.

Ottery

*A fitter fits;
A cutter cuts;
And an aircraft spotter spots;
A baby-sitter
Baby-sits —
But an otter never ots.

Though sinners sin
And thinners thin
And paper-blotters blot;
I've never yet
Had letters let
Or seen an otter ot.

A batter bats
(Or scatters scats);
A potting shed's for potting;
But no one's found
A bounder bound
Or caught an otter otting.*

From *The Biology of Algae* by Ralph Lewin. See p.500 for details.

The final third of the book is the attempt proper to debunk the notion of general intelligence as arising specifically in the school of factor analysts starting with Spearman. But by this stage the reader has been presented with sufficient examples, sufficiently carefully examined, of racial and social prejudice in the work of scientists, of distorted data, fudged analysis and twisted interpretation as to the inexpert might establish a necessary connection. Add to that the soft target of Cyril Burt, some rather inaccurate observations on the role and effects of the 11+ examination system in Britain and a remarkably detailed account of antique methods of factor analysis, and you have all the makings of a lively, plausible, opinionated and zesty potboiler.

But verbal fluency is no substitute for good arguments in the long run. The substantive discussion of the theory of intelligence stops at the stage it was in more than a quarter of a century ago. Consequently there is no account of attempts to characterize the psychological nature of general intelligence, no indication that multivariate methods have progressed beyond Thurstonian techniques, no discussion of the effects of ageing, of brain damage, of compensatory programmes, no account of modern behavioural genetics, of heritability studies other than Burt's, no hint of the current interest in cybernetic models or recent attempts by experimental cognitive psychologists to account for psychometric findings.

Gould even gives a perfectly straightforward account of what heritability would and would not mean in terms of the modifiability of intelligence, but fails to point out that such arch-hereditarians as Eysenck and Jensen have published essentially identical accounts. One is, presumably, meant to conclude that adherence to the notion that there is a measurable single dimension of intelligence necessarily involves the kind of radical nativism which was prevalent in times when genetic theory and statistical methods were in their infancy.

But this is a book with a double punch-line. Or to put it another way, Gould performs the remarkable trick of pulling the rug from under his own feet whilst appearing to stand stock still. For not only does he propose a totally unobjectionable definition of intelligence ("the ability to face problems in an unprogrammed . . . manner"), which, far from being novel, is a nice rewording of a definition proposed by Cattell in the context of a heavily factor-analytic theory and something of a commonplace amongst the intellectual heirs of Spearman, he even proposes essentially craniometric criteria of neoteny as a basis for the adaptability of *Homo sapiens*, and produces a photographic comparison of adult and infant chimpanzees with the remark that "if a picture's worth a thousand words . . .".

The truth of the matter is that Gould has

nothing to say which is both accurate and at issue when it comes to substantive or methodological points. His "fatal flaw" (the purported dependence of the notion of general intelligence on details of factor analytic technique), the unsupported assertion that "disadvantaged" groups always perform worse on IQ tests, his strictures on the application of the notion of intelligence across species boundaries, his attempt to link the use of IQ tests in Britain with a rigid class structure, all have the routine flavour of Radio Moscow news broadcasts when there really is no crisis to shout about. You have to admire the skill in presentation, but what a waste of talent. □

Steve Blinkhorn, a Senior Lecturer in Psychology at Hatfield Polytechnic, is currently working at the Neuropsychology Laboratory, Stanford University.

Beyond selfish genes

Sydney Brenner

The Extended Phenotype: The Gene as the Unit of Selection. By Richard Dawkins. Pp.307. ISBN 0-7167-1358-6. (W.H. Freeman: 1981.) £9.95, \$19.95.

THIS book, as Richard Dawkins states in his first chapter, "is a work of unabashed advocacy". It does not put forward new theories or new facts, but it advocates a special way of looking at living things and the worlds they inhabit.

This point of view, called the extended phenotype, is a development of the ideas discussed in his earlier book *The Selfish Gene* where, it may be remembered, Dawkins aimed to dispose of the organism as the unit of natural selection. He argues that this unit is the gene, or, rather, entities which are called germ line replicators, of which genes are the most important examples. Adaptations are for the benefit of these elements and not for anything else. Of course, it is admitted that these replicators are useless by themselves (you can buy them in bottles as chemicals called DNA) because selection judges them only by their effects in other entities, called vehicles, which the replicators inhabit. Biologists have given one such discrete vehicle, the organism, a special role as a unit of structure and function and it is this hold that Dawkins wants to break. He claims that replicators have extended phenotypic effects, reaching out beyond the boundaries of the organisms where they happen to be lodged into the world outside, even entering other organisms. He wants to dissolve organisms:

We see through them to the replicating fragments of DNA within, and we see the wider world as an arena in which these genetic fragments play out their tournaments of manipulative skill. Genes manipulate the world

and shape it to assist their replication.

Later he speaks of the

image of a turmoil of selfish replicators, battling for their own survival at the expense of their alleles, reaching unimpeded through individual body walls as though these walls were transparent, interacting with the world . . .

The extended phenotype is certainly a big idea and it is pressed hard in dramatic language. But does it help to view genes as active agents, manipulating, reaching, battling? I think it only obscures the essential feature of living organisms which distinguishes them from all other complex natural systems and which gives them the capacity for evolution. It is that they are propagated by passing on not themselves but an internal representation of themselves in their genes. Exactly how this correspondence is implemented is what most biologists study; knowing it must be important if only for understanding what genetic variation can generate. To a molecular biologist, all the stuff around the genes is anything but transparent; it is hard to see the replicators through the fog of complicated machinery, nor do they reach out so easily as Dawkins wants them to.

Dawkins still has to come grips with this central issue. In a section on the poverty of preformationism he would like to make a distinction between the central dogma of molecular genetics and that of embryology. The former states that information in protein cannot be translated back to nucleic acids: the equivalent in development is that "bodily form and behaviour may not be translated back into proteins". I think we agree that the former is a fact, not a question of principle, but why does Dawkins want to make the latter a principle and not simply an equivalent fact? After all, in the end we shall need to explain the effects of genes on bodily form and behaviour in terms of proteins; that is the only way for us to reach into organisms and for genes to reach out into the world. The molecular, cellular and organismic grammar connecting genes and phenotypic effects is the heart of the matter and the only way to explain how random changes in DNA can produce all the sense we observe in the living world.

Organisms exist. Even Dawkins realizes that he needs to bring back individual organisms. His extreme position has forced him into calling the organism a "phenomenon that needs explanation", and he offers only a summary sketch on the last page. He says:

It has paid replicators to behave gregariously. The phenotypic power by which they ensure their survival is in principle extended and unbounded. In practice the organism has arisen as a bounded local concentration, a shared knot of replicator power.

Sounds terrific, but what does it mean?

Sydney Brenner is a member of the scientific staff of the Medical Research Council at the Laboratory of Molecular Biology, Cambridge.

The whys (and wherefores) of behaviour

Aubrey Manning

Ethology: Its Nature and Relations with Other Sciences. By Robert A. Hinde. Pp.320. Hbk ISBN 0-19-520370-4; pbk ISBN 0-00-636237-0. (Oxford University Press/Fontana: 1982.) Hbk £9.20, \$15.95; pbk £2.95.

ETHOLOGISTS are no longer a distinctive breed; indeed the days when one could recognize them from their background and framework of theory were quite short. Large parts of the original theory built up from Lorenz's dichotomizing between the "innate" and the "acquired" in behaviour, and the energy model of drives, have been abandoned. What has endured is an approach to the animal in its natural environment which is highly productive and infectious. When Lorenz and Tinbergen built up their schools following the Second World War, ethology came as an invigorating breeze across the arid fields of so much of the white rat experimental psychology of the day. Its influence has persisted and spread widely into other areas of behavioural science.

Robert Hinde's book sets out to give an account of modern ethology and to examine this influence. He succeeds well and this is both a useful and a stimulating book. I particularly liked the brief historical introduction to each section. Since modern ethology only goes back about 30 years and Hinde himself has been a highly influential figure over most of this period, we get a direct picture of an emerging discipline as it interacts with others already established.

The plan of Hinde's opening section on "core ethology" derives from Tinbergen's classic paper of 1963, "On the Aims and Methods of Ethology". This was a landmark in the history of the subject and it represented a clear break of approach between Lorenz, who then had just published a reaffirmation of his own unchanging views on innate behaviour, and Tinbergen, who saw far more clearly that the old dichotomies (nature or nurture etc.) would no longer do. Tinbergen enunciated four types of question in relation to behaviour: those concerning its immediate causation, its development, its function (in the sense of survival value) and its evolution. Hinde has chapters covering each of these topics — he calls them ethology's four "whys". In fact they form a very good introduction which both covers some of the classic ethological studies and is also right up to date.

There is no introduction to this new *Masterguides* series itself and I am uncertain what audience they seek. Other titles include *Religion, Law and Social Anthropology*, so I imagine that "the educated layman" is once more in mind. Hinde is certainly not writing at an elementary level, nor can he bear to let

first-stage simplifications stand. Readers without some behavioural background will find parts rather heavy going and there are one or two over-complex diagrams. References to the literature abound on almost every page, and no doubt or inadequacy of the evidence is left unexposed. Just these latter qualities, of course, will commend the book to the serious student. Its main body is Hinde's assessment of the influence which ethology has had upon related fields in biology and the human social sciences. They cover a huge range, from ecology to psychiatry, and are consistently illuminating. Hinde's coverage of the literature is extraordinary and he provides good linking passages so that an overall picture of the inter-relationships of these fields is built up.

Within the biological sciences Hinde includes comparative and experimental psychology whose interactions with ethology were often stormy. The personalities, as well as the scientific achievements, of people such as Schneirla, Lehrman and Beach were of great importance in determining the course of the interchanges and Hinde gives enough of the history to make us fully aware that science is a human activity. The influence of ethology in the human social sciences has also met with opposition. The contributions it can provide are, first, the comparative method, revealing the human species as one amongst many which may share certain attributes, and secondly a secure descriptive method for the early stages of behavioural analysis. Neither of these appealed to those social scientists who emphasized the uniqueness of the human species or who concentrated solely on the cognitive and perceptual aspects of human behaviour.

Such extreme positions are rarer now. The study of non-verbal communication, of the attachment between mothers and infants and of the effects of maternal separation have all very obviously gained from ethological techniques and concepts. Hinde himself has made valuable contributions in these areas and his sharp and revealing focus is clearest in these sections of the book. The critical edge is somewhat blunted when dealing with the biological approach to anthropology where some of the wilder shores of sociobiology are dwelt upon rather too gently, but all is brought back into balance in his concluding remarks.

This book will be of great value to students and teachers in other fields who want to broaden their ethological horizons. It also provides a liberal education in these other fields for more biologically based ethologists. □

Aubrey Manning is Professor of Natural History in the University of Edinburgh.

An ode to adaptive transformation

Niles Eldredge

Darwinism Defended: A Guide to the Evolution Controversies. By Michael Ruse. Pp.376. ISBN 0-201-06273-9. (Addison Wesley:1982.) \$12.50, £8.40.

THINGS seem a bit frenetic in evolutionary biology these days. The casual observer might well think that the only theme imposing unity on this sphere of science is creationism: all evolutionary biologists agree that life has evolved, and all become willing allies in the political fight sparked by the recent successes of "scientific creationism" — which, however paltry its specific claims about natural history may be, asserts positively that life has *not* evolved. But beyond this simple and totally understandable closing of ranks, most observers see the current situation in evolutionary theory — where the object is to explain how, not if, life evolves — as bordering on total chaos. Seemingly, no two evolutionary biologists think alike these days, and we are less agreed upon things than we collectively were, say, a decade ago.

There is some substance to this view of disarray, but it has been exaggerated; we are more like a house divided than a potpourri of randomly diversified opinion. The basic problem is this:

molecular and developmental biologists, systematists and palaeontologists all think their data have something to do with evolution. This is reasonable on the face of it because, whatever else it might be, evolution involves the modification of DNA sequences, development patterns, and species and higher taxa. And the latest news from Precambrian palaeontology tells us that the process has been going on for a minimum of three-and-a-half billion years.

On the other hand we have an evolutionary theory — the "modern synthesis" — that insists that all such evolutionary phenomena are entailed by the neo-Darwinian paradigm: organisms within populations vary amongst themselves, offspring tend to resemble their parents, and because in each generation more are usually produced than can possibly survive and contribute genes to the next generation, we get deterministic, albeit statistical, patterns of change in genetic representation in succeeding generations. This, of course, is natural selection. We now feel we know more than Darwin did about why organisms resemble their parents, and what the proximate and ultimate sources of variation might be. This understanding, plus Darwin's basic arguments, combine to yield the neo-Darwinian paradigm. No less an authority than Ernst Mayr (for example in his prologue to *The Evolutionary Synthesis*; Harvard University Press, 1980), has written that the "modern synthesis" is simply this paradigm of adaptive transformation plus a single additional concept: all evolutionary phenomena, of whatever scale — from molecules to phyla — are explicable in the familiar terms of population genetics. In the apt (if by now rather trite) phrase, the theory is both necessary and sufficient to explain all known aspects of evolutionary history. However oversimplified such a characterization of the synthesis may be, it fits to a T the viewpoint espoused by Michael Ruse in his new book, *Darwinism Defended: A Guide to the Evolution Controversies*.

The simple statements of the synthesis have long been taken as very general: all sorts of phenomena which hitherto seemed to cry out for separate explanations could all now be explained under one general theory of adaptive change. And, after all, if life has had a single history, it is indeed logical to assume that there must be a single, integrated theory of the process that engendered that single history. But now the synthesis is beginning to look exceedingly narrow. For consider the effect the synthesis has had: contrary to Ruse's enthusiasm for natural selection as a "research tool", the real effect has been the relegation of molecular and

developmental biology, systematics and palaeontology to the role of mere reportage. Evolutionarily-inclined members of these diverse disciplines have been told, in effect, to sit back, relax and simply recount the results of the evolutionary process — and leave it to the population geneticists to explain how it all happened. From time to time, members of one or more of these various satellite evolutionary disciplines have objected, insisting once again that their data — their molecules and fossils — must have something more direct to say about the validity of this or that notion about how life really does evolve. Now their voices are becoming more strident.

Michael Ruse is aware of this state of affairs. Hence the book, which has two titles, accurately conveys his priorities. First, it is a stout defence of what Ruse calls Darwinism. Secondly, and only through the glasses of a truly arch-conservative, the book touches upon virtually all the major issues of which I am aware. I call Ruse an arch-conservative advisedly: to this staunch *defensor fidei*, genetic drift is still a heresy. Anything smacking of the stochastic, or even of "neutrality", has him shaking his head and wagging his finger in a trice.

Darwinism, to Ruse, seems to be adaptation through natural selection. I say "seems" simply because his enthusiasm sometimes gets the better of him, and Darwinism in places comes close to meaning simply "evolution". Ruse knows better, of course, and generally makes the distinction between the notion that all living things are interrelated, and specific ideas (of which Darwinism is one) about how life has evolved, satisfactorily explicit.

Ruse has structured his book rather well. There are five main sections; the first three deal with his view of Darwinism yesterday, today and tomorrow. The first section gives us a brisk account of the man and his main works — with emphasis on Ruse's analysis of the *Origin* as a three-part essay. "Darwinism Today" tells us of the "coming of Mendelian genetics", what it is and what it means. Part 3, "Darwinism Tomorrow", is an eclectic grab-bag of topics, ranging from the "origin of life" to the "challenge from palaeontology". By this stage the reader will have no difficulty in predicting Ruse's line of argument. Those areas of contemporary biology stressing adaptive explanations through selection are "good", while he takes a dim view of disciplines and ideas which seem, somehow, to diminish the central role of adaptation. Thus sociobiology is "good", but punctuated equilibrium theory has the malodorous air of yesterday's half-eaten fish and chips. Part 4, "Darwin and Humankind", deals with the notion that *Homo sapiens* has had an evolutionary history, proclaims human sociobiology as on the side of the angels (if only metaphorically) and says that one should not adduce nasty ethics of the dog-eat-dog

FUSARIUM:

Diseases, Biology,
& Taxonomy

ed by P E Nelson,
T A Toussoun, &
R J Cook

This volume brings together what is currently known about the pathology, ecology, genetics, cytology, physiology, biology, and taxonomy of this important fungus genus.

Penn State University Press,
April, 560 pp, 7" x 10",
65 illus, £27.65

NATURAL HERITAGE

Classification,
Inventory & Information

by Albert E Radford,
Deborah K Otte, Lee J Otte,
Jimmy R Massey, Paul Whitson, &
contributors

The ultimate goals of this study are to foster the conservation of natural heritage, to advance field training in ecology, systematics, and evolution, and to determine site-population relationships for species of special concern.

University of North
Carolina Press,
485 pp, 8½ x 11",
maps & figures, £17.50

AUPG

1 Gower Street
London WC1

variety just because sociobiology can explain so much of our behavioural history. I found this section, and the two-chapter, fifth and final section on creationism, unprepossessing.

The guts of the book lie, appropriately enough, in its middle, in parts two and three. Here Ruse tells us what evolutionary biology really is, and should be, all about. I find it a dismal picture. The approach Ruse advocates is pure storytelling — or, as he himself puts it, continued indulgence in evolutionary biology's "favorite parlor game": concocting plausible scenarios of how the elephant got its trunk, the rhino its wrinkled skin and the giraffe its long neck. To Ruse, we *know* the mechanism, and it's all just a bunch of clever games applying this mechanism of genetic change to explain literally all manner of evolutionary phenomena.

Most critics of contemporary evolutionary theory, myself included, would agree that natural selection is a deterministic mechanism accounting for generation-by-generation change in gene frequencies within populations. It is even testable, given the appropriate data — which involve gene frequencies and generations. But reducing, say, ammonite evolution to a just-so story with natural selection as the hero is, as Ruse says, nothing but a parlour game. The difference between us is that Ruse thinks this state of affairs is just fine. I think parlour games are fine too — but I still cherish the fantasy that I am doing, or at least am trying to do, science when I'm practising my profession.

Ruse uses analogy and circumstantial evidence as his cornerstone criteria for establishing the validity of scientific ideas. Claiming Darwin used Herschel and Whewell as models, and that analogy and inference were what made science "good" in the 1850s, Ruse (as he does so much throughout the book) in effect argues that what was good enough for Darwin should suffice for us as well. There is little here of the spirit of enquiry. We are not enjoined, for example, to hold our notions lightly, to walk humbly before Mother Nature, ready to modify our schemes should the evidence of our senses not rhyme with our fondest notions.

Nowhere in this single-minded tract is Ruse's dedication to the almighty principle of natural selection better displayed than in his little section on trilobite vision. The mathematically "perfect" shape of phacopid and dalmanitid lenses is, to Ruse, exquisite evidence of the power of natural selection. I confess I cannot fathom the difference between Ruse's argument and the older creationist argument from design: see this organ system; observe its intricacy! Only (God, natural selection) could have fashioned such a marvellous organic machine! There is a difference, of course: God, as a supernatural being, does not belong in science, whereas natural selection patently does. But used in this inappropriate fashion, natural selection

becomes a mere substitute for the Creator. It tells us nothing, really, about trilobite eyes or anything specific or meaningful about how they came into existence.

In short, Ruse is a reductionist — all large-scale evolutionary phenomena are readily explicable in terms of population genetics. But there is more than mere reductionism here: he also dismisses molecular biology, simply because population genetics revealed the contents of Darwin's "black box" of heredity, and molecular biology, Ruse avers, has not changed *that* one whit. All remains secure. That there are patterns of organization with historical implications emerging from molecular and developmental biology that seem to require their own mechanisms *in addition* to those of population genetics is missing in Ruse's world view. Indeed, in his reflexively negative reaction to the winds of change, Ruse fails to see that the newer arguments are nearly all additive: natural selection explains changes in gene content and frequency within populations, but more than populations are involved in evolution. Another example: Ruse doesn't even bother to mention the idea that species are individuals — the ontological claim underlying the view that macroevolution

may be something more than just scaled-up microevolution.

Darwin and his ideas need no defence. I agree with Ruse: Darwin was a good scientist. In that spirit I would rather imagine Darwin enjoying the spectacle of biologists vigorously debating his, and descendant, notions over a century after the *Origin*. Ruse, appropriating Darwin's shade as his alone, seems to think that to question some of Darwin's notions is somehow to impugn the man. So he springs to the defence. It is a clumsy defence, one that certainly does Darwin no credit. As a statement of the nature and effective extent of neo-Darwinism, it is nowhere as good as, say, the September 1978 issue of *Scientific American*, or any of a number of recent texts. As a book aimed at a general audience, it projects a jolly *Alice in Wonderland* sort of picture that, I fear, will not do overly much for our collective image in the long run. Darwinism, indeed, the entire field of contemporary evolutionary biology, deserves far better. □

Niles Eldredge is at the American Museum of Natural History, New York. His most recent book, *The Monkey Business: A Scientist Looks at Creationism*, will shortly be published by Washington Square Press.

A new paradigm for evolutionary change?

R.D. Martin

The New Evolutionary Timetable: Fossils, Genes and the Origin of Species. By S.M. Stanley. Pp.222. US ISBN 0-465-05013-1; UK ISBN 0-633-7022-8. (Basic Books/Harper & Row: 1982.) \$16.75, £9.50.

IT IS, perhaps, appropriate that the centenary of Charles Darwin's death should be approximately marked by renewed controversy over the process of evolutionary change. The latest debate concerns one of the most fundamental features of Darwin's theory of evolution by natural selection — *gradual* change within individual species populations over time. An alternative interpretation, stated with particular clarity by Eldredge and Gould in a seminal paper published in 1972¹, is that most evolutionary change occurs when a new species develops (during a "speciation event") and that there is little change subsequent to establishment of a species.

In some circles, at least, replacement of the Darwinian concept of "gradualism" by the new model of evolution by "punctuation" has been hailed as a fully-fledged scientific revolution (a Kuhnian paradigm change). To date, however, the debate has raged largely among academics, and for this reason Steven Stanley's *The New Evolutionary Timetable* is both timely and

welcome. Following hard on the heels of his influential scientific text *Macroevolution: Pattern and Process* (W.H. Freeman, 1979) this new book attempts to present the case in terms understandable to the non-specialist. Stanley achieves his aim with considerable success, carrying the reader along with a style that is at once lively and informative. Penetrating insights abound as the arguments are presented, and frequent touches of wry humour add a special touch. (Any modern biologist struggling for a command of the literature will surely be relieved to learn that Darwin actually possessed an uncut copy of Mendel's monograph on hereditary mechanisms!) Certainly, a reading of this book leaves one with plenty of food for thought, regardless of the correctness of Stanley's own conclusions. This is a particularly productive time for the remodeling of evolutionary theory, and *The New Evolutionary Timetable* conveys the excitement felt by many biologists involved.

Underlying the concept of punctuated evolution is an undeniable fact derived from recent palaeontological studies: numerous fossil species have been found which exhibit a remarkable degree of stability ("morphological stasis") over long periods of geological time. It must at once be added that such stasis can only be recorded for certain body parts (molar teeth of mammals or shells of molluscs, for

¹Eldredge, N. & Gould, S.J. in *Models in Paleobiology* (ed. Schopf, T.J.M.), 82-115 (Freeman & Cooper, 1972).

example) and it is reductionist to conclude that entire organisms remain unchanged over time. Differential evolution of individual body components (mosaic evolution) is a widely recognized phenomenon and there is, for instance, no indication of stasis at the level of the structural gene. Stanley makes little mention of the substantial literature on molecular evolution, which suggests a hiatus between rates of evolution of single proteins and those of whole organisms. It is all too easy to slip into equating fossilized skeletal parts with entire organisms, and this slip is a recurrent feature of the literature of punctuated evolution. Stanley, for example, asserts that *Homo sapiens* has undergone almost no bodily change since appearing in Europe some 40,000 years ago.

The second aspect of the punctuational model is rapid evolution during speciation, which is generally assumed to occur in relatively or very small, peripheral populations. Such rapid evolution must, however, be *inferred* from the fossil record and involves the assumption that the fossil record is sufficiently complete to document this process unequivocally. In a few special cases this may be true, but the fossil record as a whole is pitifully inadequate — there is, for example, only one major fossil site recording the occurrence of mammals in the entire African continent for the first 35 million years of the Tertiary. We really

need to have estimates of the relationship between the likely catchment areas of relevant fossil sites and the geographical ranges of equivalent modern species; in the absence of such estimates a reasonable guess is that fossil sampling is in fact extremely patchy.

Nevertheless, the picture now emerging from the fossil record is unsettling, and our ideas of evolutionary processes must undoubtedly progress to account for long-term stasis, at least in individual body parts, in species populations. The question is whether such progress requires no more than healthy readjustment of basic theory or whether the changes which occur during speciation dictate recognition of a radically new process ("quantum speciation" or "macroevolution"). Stanley is convinced that the latter is the case. Perhaps understandably, he does not fix upon a single basic mechanism of "macroevolution", though he does somewhat vaguely implicate chromosomal reshuffling. Everything really hangs on the rapidity with which speciation takes place. Stanley states that he is prepared to envisage thousands of generations as the time-scale for "quantum speciation", and this could well be accommodated within a traditional model of evolutionary change. Stanley feels that the fossil evidence for human evolution provides us with one of the best illustrations of punctuation, yet "thousands of generations" could occupy a considerable proportion of the average geological duration of the hominid species currently recognized. In any case, the situation is far more complex than Stanley suggests, notably in the existence of intermediate forms (see Cronin *et al.*, *Nature* 292, 113–122). Stanley himself is somewhat unclear about the actual limits of his model as well. He suggests (p.157) that human races ("subspecies") may have arisen by punctuated evolution. Quite apart from the fact that here he is treading on extremely contentious ground in sociopolitical terms, it is difficult to see what "quantum" differences might define the human races in any context relevant to punctuated evolution.

In the absence of a clear model of the mechanism of "quantum speciation", the case for a radically new paradigm remains somewhat unconvincing, as was the case with continental drift prior to the recognition of sea-floor spreading. Stanley devotes a whole chapter to demonstrating that his doubts about conventional theory do not open the door to creationism; but without a clear concept of the mechanism involved in "punctuated evolution" he is unlikely to convince any creationist! Stanley is right to alert the general reader to the surprising new findings of palaeontologists, but he really goes much farther than the present evidence permits. □

R.D. Martin is Reader in Physical Anthropology at University College London and Visiting Professor in Zoology at Birkbeck College, University of London.

Two sides of Darwin

A.J. Cain

Darwin. By Jonathan Howard. Pp.101. Hbk ISBN 0-19-287557-4; pbk ISBN 0-19-287556-6. (Oxford University Press: 1982.) Hbk £5.50; pbk £1.25. To be published in the USA later this year.
Darwin. By Wilma George. Pp.160. ISBN 0-00-636502-7. (Fontana: 1982.) £1.75.

It is a sign of Darwin's relevance that Jonathan Howard's book is in the OUP *Past Masters* series, Wilma George's in the Fontana *Modern Masters*. They are nearly identical in their rather nasty format (dirty-coloured paper, undistinguished type-face and over-gorgeous covers, rather like the trash-novels on a railway bookstall), but do not be put off by first appearances — their intellectual content is very good, and both are well worth the money. Both are written by practising zoologists — a refreshing change from so many books on Darwin — but zoologists of very different interests. Jonathan Howard is a Principal Scientific Officer at the ARC Institute of Animal Physiology at Babraham, Wilma George is a University Lecturer in the Department of Zoology, Oxford. There is also a useful difference of sex; some of Wilma George's comments on the acceptability of Darwin's theories of sexual selection and male aggression to Victorian men are particularly enlightening.

Jonathan Howard remarks in his preface that

With the centenary of Darwin's death comes a widespread mood of scepticism and unease about the validity and significance of Darwin's contribution to knowledge. The time is certainly right to try to describe Darwin's scientific work briefly and in plain language.

While fully recognizing the importance of Darwin's work for biology, he says "I hope the book is fair in pointing out aspects of Darwin's thinking which lack consistency or have failed to stand up to critical scrutiny". And he rightly emphasizes that whatever others have made of Darwinism in philosophy or sociology, and however true it is that "human life and human society are to a certain extent biological issues", it is Darwin's contribution to biology on which he must stand or fall. "Darwinian philosophy of Darwinian society are post-hoc constructs that had no place in Darwin's thought".

The chapters take us successively through Darwin's life; the foundations of Darwinism; natural selection and the origin of species; the evidence for evolution by natural selection; sex, variation and heredity; man; perfection and progress; Darwinism and ideology; and Darwin as a scientist — an evaluation. The style is mostly good and the exposition clear. Just occasionally it would have benefited from critical reading — for example I can make no sense out of Howard's comment on

Just Published!

EARLY LIFE

Lynn Margulis
Boston University

"... a fascinating account of biochemical, metabolic, and physiological evolution during the age of microbes — the first 3 billion years of the history of life on earth... clear, accurate, and a pleasure to read."

James C. G. Walker
University of Michigan

"Great resource and beginning reading for the non-specialist who wants to learn about the evolution of early life."

Kenneth H. Nealson
Scripps Institution of Oceanography

Feb. 1982, 176 pp., 62 illus. Glossary.
\$9.95 (paper), \$16.50 (case). *Science Books International*, 51 Sleeper Street, Boston, Mass. 02210. Write to Jane Wescott for purchase or examination copy, or call (617) 426-2224.



Darwin's rejection of Christianity (p.9):

Darwin reserved this uncompromising expression of his religious views for the private autobiography that he wrote for his family. Had he not done so it is inconceivable that T.H. Huxley could have abandoned his war against hypocrisy and helped to carry Darwin's coffin into Westminster Abbey in 1882.

It is inconceivable that Huxley did not know Darwin's views on religion — Darwin was not reticent about them to correspondents — and that he should not have applauded them. Where was the hypocrisy?

Howard's book is essentially a careful exposition and critique of Darwin's arguments in the *Origin* (and in the *Descent of Man* for the evolution of man). It is in the structure of the arguments that he is principally interested. Of course he does not neglect the factual evidence, but produces usually only the minimum of it — for example his comments on the cogency of geological, geographical, taxonomic and embryological information in establishing that evolution had occurred (pp.37–49). Similarly, in summarizing the modern position on natural selection he merely remarks (p.87),

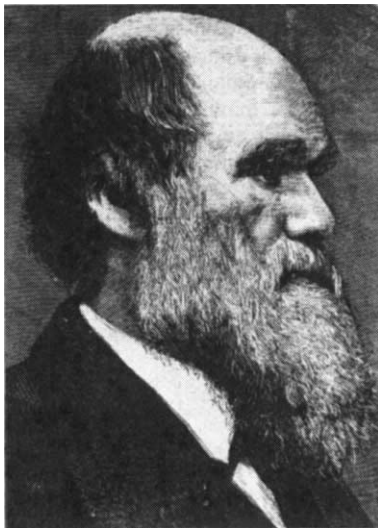
Natural selection and the evolution of natural populations have been observed directly, even in man. There is now no doubt that natural selection is a mechanism of evolutionary change. Whether it is the exclusive mechanism of evolutionary change is still occasionally debated.

And he accepts Darwin's own belief that he had spent too much time on barnacles (p.5) without analysing what he had got from them.

Howard's exposition of ideas and their inter-relatedness is much more full, and he

seems especially interested in predictability and unpredictability and their role in the evolutionary debate. He remarks (p.24):

If mutual competition between organisms was a more important selective influence than mere climatic or geographic change, then the direction of evolution of a species or of an interacting group of species, became effectively impossible to determine.



And, after a supporting quotation from the *Origin*,

It is important to recognise this instance for what it is, namely a hypothetical scenario for fluctuations in frequency of different kinds of organisms that are all subtly related to each other, initiated by introducing into a dynamic equilibrium a single small change . . .

Here is the physiologist pre-adapted for seizing the complexities of a dynamic system and comprehending the modes (and pitfalls) of its analysis — a valuable approach.

Wilma George's book, written in a particularly attractive and lively style, inevitably covers much of the same ground but is effectively complementary to Howard's in its mode of treatment. She has travelled extensively, realizing the facts of natural history, adaptation and biogeography at first hand — an equally valuable approach. She uses all of Darwin's works very effectively, showing how his interests grew on a wide front, and provides the best short account that I know of just why he was so interested in barnacles, what he got from them and how the ideas acquired led on to later interests (polymorphism, sexuality, insectivorous plants and the fertilization of orchids, for example). Her insight into the development of his ideas on coral reefs in relation to other geological ideas is particularly clearly expressed, and she relates Darwin most effectively to other workers before, at the time, and after. She gives a more concrete exposition of subsequent work in population genetics, natural selection and speciation, than does Howard.

Either book can be read with pleasure by itself; the two combined make a first-class introduction to the subject. Both authors make suggestions for further reading, to which should be added Dorothy Nelkin's spine-chiller *Science Textbook Controversies and the Politics of Equal Time*, published by MIT Press in 1977. Both glance at the present situation in evolutionary thought, and both expatiate on the forces of unreason, but neither, it seems to me, appreciates just how powerful and dangerous they are. Nelkin's book is a necessary addition. □

A.J. Cain is Derby Professor of Zoology at the University of Liverpool.

One foot in sea, and one on shore . . .

D.R. Newth

The Aquatic Ape: A Theory of Human Evolution. By Elaine Morgan. Pp.168. ISBN 0-285-62509-8. (Souvenir Press, London: 1982.) £7.95. To be published in the USA later this year by Stein & Day.

*Sigh no more, ladies, sigh no more,
Men were deceivers ever;
One foot in sea, and one on shore,
To one thing constant never.*

ONE of the few changes made by Darwin in preparing the second edition of the *Origin of Species* was the exclusion, in obedience to Lyell, of any reference to the "secondary whale". The evolution of aquatic mammals still presents troubling problems, but at least, for the most part, those mammals taking to water do not appear to have had second thoughts. It was, therefore, brave of Alister Hardy to

suggest, nearly a quarter of a century ago, that mankind had passed through a semi-aquatic phase and that some of our less easily explained biological features were its legacy.

Hardy laid stress upon our relative hairlessness — and upon the orientation of the hairs that we do possess, upon bipedalism, upon subcutaneous fat, upon our ability to swim, and upon the form and sensitivity of the human hand. All seemed to make sense as adaptations to littoral life, less aquatic than that of present-day otters, but one sufficiently dependent upon shallow waters to make the transition to bipedal locomotion easier to envisage in view of the buoyancy enjoyed by the wading anthropoid. Although Hardy did not attract much support for his idea he was content to wait for the matter to be decided by the fossil precursors of the australopithecines when their discovery should fill the late

Miocene-early Pliocene gap in the story.

In bringing us, with Sir Alister's blessing, an extended statement of the case for his ideas, Mrs Morgan adds arguments of her own. She sees in face-to-face copulation an important behaviour pattern shared by mankind and many aquatic mammals, but not seen in infra-human primates. Emotional weeping is another such behaviour. She also believes that speech as a mode of communication may have evolved as the limitations of smell and sight in an aquatic environment made themselves felt. In this, and in the suggestion that underwater childbirth may once have been the norm, she seems to envisage a more completely aquatic phase than Hardy originally did.

The additional arguments do not, I think, greatly strengthen the case. Mrs Morgan herself sees face-to-face copulation in mankind as a consequence of bipedalism and hence, at most, as a secondary consequence of an aquatic phase. On the shedding of tears the position is not simple. While seals and sea

otters apparently weep for emotional reasons as do we, so too do Indian elephants. Elephants like whales have, however, reduced their lachrymal gland and increased the Harderian gland contribution to tears. Mrs Morgan is willing to accept elephants as fellow travellers to the water and back, on this and other grounds, but until we are more confident about the functions of copious tear production it seems best to be cautious.

However the fossil record will doubtless have the final say, and it is becoming yearly more complete. Mrs Morgan re-publishes the suggestion of another supporter of the Hardy hypothesis, L.P. La Lumiere. He believes that the aquatic phase was experienced during early Pliocene times on the coast of the Red Sea. At a time of presumed deforestation, changes in land

and water levels providentially provided an island refuge. La Lumiere offers precise predictions of where the missing links will be found.

This book is written for a lay public and is too brief, and too superficial, to satisfy a serious critic. The bibliography, for example, fails to include many of the authorities referred to in the text. The book is also marred by occasional carelessness — I had to read one passage several times to extract from it a meaning *other* than that women and sirenians were alone in having two pectoral mammary glands. But the writing is lively and enthusiastic and it will surely engage the interest of those for whom it was written. □

D.R. Newth was formerly Regius Professor of Zoology in the University of Glasgow.

... I expected a lot from my associates — hard disciplined work and the ability to accept my criticisms . . . I have tried to be fair, honest and helpful, and not to demand more of others than I do of myself . . . We criticised each other ruthlessly — we knew it was done honestly, in good faith and in a spirit of helpfulness.

Few Englishmen would write such solemn stuff about themselves. I cannot imagine Francis Crick declaring in his memoirs that he demolished my cherished 1949 model of haemoglobin, "honestly, in good faith and in a spirit of helpfulness", rather than admitting to a certain mischievous satisfaction at the dastardly deed.

For people who complain that money for research has become hard to come by, Krebs's early career makes salutary reading. His father, a doctor in Hildesheim, had to support him throughout his medical studies. After graduating Krebs wanted to do research, but paid research posts were rare, never advertised and obtainable only by what used to be known in my native Vienna as *Protektion*, which signified a mix of sycophancy, nepotism and the old boy network. Eventually Krebs found an unpaid post at the Third Medical Clinic in Berlin where another biochemist who later made his name, Bruno Mendel, became one of his colleagues. One night the Mendels were invited to the Einsteins when Otto Warburg was among the guests. When Warburg told Mendel that he was looking for a collaborator, Mendel recommended Krebs and also raised the money privately to pay him a modest salary. Krebs regards his four years with Warburg as formative for his career and expresses the greatest admiration for Warburg's scientific genius, regardless of his despotic, egotistical and sometimes malicious behaviour, his lack of confidence in Krebs's own talent and his refusal to help him find a university post where Krebs would have to attach himself "to some old ass of a professor". Only in medicine would Krebs be able to make a living, Warburg told him. Warburg demanded of his staff that they work punctually from eight to six, six days a week, a precept which Krebs himself seems to have followed for much of his life. At 11 o'clock on the Monday morning after reading of Warburg's Draconian regime, I walked into my own biochemistry laboratory to find that only one of my three collaborators had arrived yet. "The other two were still working when I left at midnight", the first reported. This convinced me again that the free coming and going of Cambridge is more conducive to dedicated research than the iron discipline of old Berlin.

Krebs made his first great discovery, the ornithine cycle of urea synthesis, in 1931 and 1932. He used the tissue slice and manometric techniques which Warburg had taught him; another decisive factor was his own invention of the Krebs-Ringer

A cycle ride to Stockholm

M.F. Perutz

Reminiscences and Reflections. By Hans Krebs in collaboration with Anne Martin. Pp.250. ISBN 0-19-854702-1. (Oxford University Press: 1982.) £12.50, \$19.95.

ON 14 December 1932 the dean of the medical faculty of the University of Freiburg, Professor E. Rehn, reported to the Ministry of Education,

As an assistant physician Dr Krebs has shown not only outstanding scientific ability, but also unusual human qualities. His . . . paper on the synthesis of urea in the animal body . . . will be regarded as one of the classics of medical research .

Four months later, shortly after Hitler seized power, this same dean sent him a "Notification of Immediate Removal from Office", obediently implementing the Minister's orders against members of the Jewish race.

Krebs kept several such letters and also various press cuttings from those days and had them reproduced as facsimiles in his book. There is the "Manifesto Against the UN-German Spirit" published by the Freiburg Student Union and posted in the University: "Our most dangerous adversary is the Jew and he who serves him. The Jew can think only as a Jew. If he writes German he lies. The German who writes German, but thinks un-German is a traitor . . .". The Rector of the University, von Möllendorf, ordered the posters to be removed, whereupon the Minister replaced him by the existentialist philosopher Martin Heidegger who endorsed them and proclaimed, "German students! . . . Your existence must not be ruled by learned axioms and ideas . . . Only the Führer himself constitutes today's and tomorrow's reality and law. Daily and hourly must you fortify your faithful

obedience . . .". Most of Krebs's colleagues acquiesced, but one Dr Arthur Jores dared send his and Krebs's former Jewish chief, by then in New York, a reprint with a personal dedication. The envelope was opened by a Nazi colleague who denounced him. Jores was dismissed from his post, publicly branded as an enemy of his country and, in due course, imprisoned.

Despite these harrowing experiences and the extermination of 20 of his relatives Krebs does not follow the example of Einstein who refused ever to set foot in Germany again, but declares in the book that "to be anti-German seems to me just as bad as being anti-semitic". After the War he therefore persuaded the Biochemical Society to re-establish contact with Germany by inviting several known anti-Nazis to the First International Congress of Biochemistry at Cambridge.

Krebs's discovery of the ornithine cycle of urea synthesis had caught the eye of F. Gowland Hopkins, then Professor of Biochemistry at Cambridge and President of the Royal Society. When he heard of Krebs's dismissal, he immediately invited him to Cambridge. The Rockefeller Foundation which had already supported Krebs's research at Freiburg, provided the money, and continued to support his research for the next 30 years. Krebs was deeply touched by the warmth of his reception at the Biochemistry Department at Cambridge and by the generous hospitality and friendliness of English people generally. His book is permeated by affection for England, which he epitomizes by Carl Zuckmayer's, another refugee's, saying that "home is not where a man is born but where he wants to die".

How Prussian Krebs remained all the same! He writes,

solution. I found it exciting to read of his groping in the dark, methodically testing all conceivable intermediates until he discovered that "one molecule of ornithine could bring about the formation of more than twenty molecules of urea, provided that ammonia was present". From that moment the tracing of the other intermediates followed logical steps. It was the first biological process in which the intermediates were found to play a purely catalytic role. Krebs made this fundamental discovery while in charge of a medical ward with over 40 beds, which makes his feat even more remarkable.

Krebs's unravelling of the citric acid cycle in 1937 was to win him even greater fame, but his letter to *Nature* announcing it was rejected by the editor, Sir Richard Gregory, who at that time took it upon himself to judge the scientific worth of most of the communications sent to him. (It is not known whether Krebs's paper was refereed, but that seems unlikely — Ed.)

Objections were raised at first against both the cycles. Those against the ornithine cycle were later found to be based either on wrong experiments or incorrect interpretation, while those against the precise chemistry of the citric acid cycle seemed fundamental. Biochemists argued that radioactively labelled CO_2 introduced into the cycle should become randomly distributed among the two carboxyl groups of α -ketoglutaric acid, an intermediate two steps after citric acid, because an enzyme would not be able to distinguish between the two symmetry-related carboxyl groups of citric acid. In fact, only the carboxyl nearest to the keto group of α -ketoglutaric acid was found to be labelled; hence, it was concluded, citric acid could not be an intermediate in the Krebs cycle. That was in 1941. I would have been desperate if an apparently valid objection had been made to my most fundamental discovery in which I could detect no flaw, but Krebs writes as though it had never cost him any sleep. Was he really so placid that he did not continuously turn over in his mind all conceivable explanations of the paradox, or did the sunshine of his later glory dissolve the memory of the seven clouded years that were to elapse before A. Ogston, in a brief and classic note to *Nature*, pointed out the fallacy in the objection: a symmetric molecule attaching itself to an enzyme at three points may give rise to only one of two possible asymmetrical reaction products. This was the birth of the concept of prochirality.

The glory was first heralded in October 1952 by eager journalists who told Krebs that he would shortly be awarded the Nobel Prize, but the rumours proved false and S.W. Waksman received it instead. Krebs relates proudly that the rumours left him and his wife unruffled. But did they really? Nine years later similar rumours about John Kendrew and me were floating

around our laboratory. We doubted them until my secretary rushed in, flourishing two telegrams, one addressed to Kendrew and the other to me. This was it. When we had eagerly torn them open we found them to be from the Pontifical Academy in Rome, enquiring how many reprints we wanted of the papers we had read there the previous autumn. We also pretended to a stoic calm. Unruffled or not, Krebs did receive the Prize the following year.

After nearly 20 happy years and the founding of a flourishing school of biochemistry at Sheffield, Krebs became

Professor of Biochemistry at Oxford and stayed there until his death last November. I was surprised by his statement that Oxford had remained in the forefront of learning for 600 years. Had he never read Edward Gibbons's description of its decline into sloth in the eighteenth century when dons could not even be bothered to teach their students? The picture of Oxford University in Krebs's own time conjured up in the book is of a citadel of unjust privileges jealously guarded by the Party members. Krebs got himself elected to the Central Committee by the under-

Scherzo metabolico cantabile

Thomas A. Scott

The Biochemists' Songbook. By Harold Baum. Pp.62. Pbk ISBN 0-08-027370-X. (Pergamon: 1982.) £2.45, \$4.95.

How refreshing, in the 1980s, to find a professor of biochemistry writing, not a solemn research text, but humorous biochemical verses. Thirteen important topics of biochemistry have been rendered by Professor Baum into light-hearted songs and set to appropriate popular tunes with piano accompaniment. The biochemical accuracy of the outrageous doggerel can be checked at a glance, because each song appears with the conventional scheme or account.

According to the author, all the songs were written while travelling on the top deck of a No. 22 bus; this I do not believe. He also points out that they are intended for communal singing, ideally with a blood alcohol level of 35 mg per cent; this I do believe. Some of the scansion is tortuous and not always immediately obvious. While the alcohol will not make them any easier, it will stop you worrying whether you get them right.

Perhaps the greatest challenge is in the first verse of "Protein Biosynthesis":

*Introns and exons, changes post-transcriptional, and all
Glycosylations, don't alter such basics at all*

to the tune of My Bonny Lies Over the Ocean. In contrast other compositions have a reassuring and earthy quality, for instance:

*If you gobble tagliatelli,
Chicken soup with vermicelli,
You'll acquire a sagging belly —*

are the opening lines of "Fatty Acid Biosynthesis" set to Men of Harlech.

Surprisingly, the author has failed to exploit the rousing coda of Macnamara's Band in "The Pentose Phosphate Shunt". Admittedly, this loud and rasping interlude is normally left to the trombones, but it is tailor-made for Professor Baum's particular style of ethanolic chorus; in the second edition perhaps? Also, these days, when so many bar pianists are over forty, and all the young wags play the guitar, this already excellent songbook might be made more usable by the younger generation by the inclusion of harmony frets or chord names for each tune.

The Biochemists' Songbook is dedicated jointly to Sir Hans Krebs on his eightieth birthday and to the author's wife. Sir Hans has also contributed a foreword, and there is a certain poignancy in the publication of the *Songbook* so shortly after his death. His cycle is set to Waltzing Matilda, and this year the International Congress of Biochemistry will be held in Australia, where it is no novelty, indeed it is a long established tradition, to sing this tune communally and with a blood alcohol level rather in excess of that recommended by Professor Baum.

Thomas Scott is Senior Lecturer in Biochemistry at Leeds University.

privileged Solidarity of Active Researchers, but all his brave attempts to reform The System were defeated by the strength and cunning of the Old Guard, who even went so far as to refuse money offered by the University Grants Committee for the creation of additional science professorships, lest this should lead to the appointment of more Dissidents and Enemies of the Party like Krebs. Despite these rebuffs, Krebs describes his life at Oxford as a happy and successful one. Asked once for a guiding motto, Krebs replied "Never put off till tomorrow what you can do today", which sounds like one of those recipes for virtue Victorian children were made to embroider and hang over their beds. The King of France says it less prosaically in *All's Well that Ends Well*:

Let's take the instant by the forward top;
For we are old, and on our quick'st decrees,
The inaudible and noiseless foot of time
Steals ere we can effect them.

Krebs emerges from this book as a dynamic scientist and an engaging, warm-hearted individual utterly devoted to his research and teaching. He quotes Noel Coward's saying that "work is fun, there is no fun like work". I agree. □

M.F. Perutz is a Member of the Medical Research Council Laboratory of Molecular Biology, Cambridge. With John Kendrew he was winner of the Nobel Prize for Chemistry in 1962.



AN ENUMERATION OF THE FLOWERING PLANTS OF NEPAL

Vol.1 Gymnospermae & Monocotyledones.
H. Hara, W.T. Stearn & L.H.J. Williams.
June 1978. **£22.50**

Vol.2 Dicotyledones I.
H. Hara & L.H.J. Williams.
June 1979. **£22.50**

Vol.3 Dicotyledones II.
H. Hara, A.O. Chater & L.H.J. Williams.
May 1982. **£32.00**

The publication of the third volume completes this joint project of the British Museum (Natural History) and the University of Tokyo. It is based on the extensive botanical exploration of Nepal by British and Japanese scientists during the last 25 years.

**Publications Sales,
British Museum (Natural History),
Cromwell Road,
London SW7 5BD**

Circle No.31 on Reader Enquiry Card.

Professors, parasites and public health

Anthony C. Allison

New Guinea Tapeworms and Jewish Grandmothers: Tales of Parasites and People. By Robert S. Desowitz. Pp.224. ISBN 0-393-01474-6. (W.W. Norton:1981.) \$12.95, £8.75.

DURING the period following the Second World War, a few pioneers of a particular type flourished. They were motivated by curiosity about medical problems in primitive societies, the people themselves and how social organization affects health. In an era of liberal funding they were able to spend a great deal of time travelling to New Guinea, South-East Asia, the Australian outback, East and West Africa and other appealing places, collecting thousands of samples of blood, brains and other materials. For some time they were regarded as amusing, if somewhat irresponsible, amateurs, who should settle down and do serious research in Bethesda or London. But they enjoyed themselves, satisfied their curiosity and carried out research of real quality, sometimes solving problems that had for decades defied frontal attack in the United States and Western Europe.

Carleton Gajdusek transmitted kuru, a degenerative disease of the central nervous system, from cannibals of the New Guinea highlands to chimpanzees. His energy and vision played a large part in development of the concept of slow viruses. Baruch Blumberg found that certain sera of North Americans who had received blood transfusions precipitated a component in the sera of some Australian aborigines. Out of this seemingly trivial observation came the identification of hepatitis B virus and the development of a test that has nearly eliminated post-transfusion hepatitis in the United States and Western Europe. Both of these developments won Nobel prizes. Following Denis Burkitt's observations on the prevalence of a particular type of lymphoma in Uganda and other East African countries, Tony Epstein identified in the lymphoma cells the herpesvirus that bears his name. This was found to transform human lymphocytes, to cause infectious mononucleosis and to be associated with a type of cancer of the nasopharynx common in parts of East Africa and China. It is the first likely candidate for a human cancer-inducing virus. The identification of the single amino-acid substitution in sickle-cell haemoglobin, and the demonstration that this gene confers resistance to malaria, were notable contributions to molecular and population genetics.

Included in this group of individualists is Bob Desowitz. After an orthodox training in tropical medicine, he worked on sleeping sickness in Nigeria and on various parasitic diseases in Singapore. He then became Professor of Tropical Medicine in Hawaii,

and has travelled as a consultant for the World Health Organization and national governments to New Guinea, Indonesia, Burma and other places. This is a life that many might envy, and its highlights are recounted in this book of essays. Although the title might seem flippant, the book is instructive and entertaining.

Desowitz has a clear and readable prose style. Most medical students and biologists are bored by the life cycles of parasites. It helps to know that the fish tapeworm *Diphyllobothrium latum* was introduced into the fish of the lake region of Minnesota and Wisconsin by Scandinavian fishermen and was not killed during the delicate cooking of gefilte fish by New York grandmothers. The book is packed with easily assimilated information about parasites. Desowitz is an amusing raconteur and enjoys the detective side of public health, such as finding that a common brain infection among the Ekari of New Guinea was cysticercosis (tapeworm disease) transmitted by pigs given to the people by President Suharto of Indonesia to sweeten the military occupation.

The two central themes of the book are, first, that man-made ecological-environmental changes have been responsible for perpetuating and intensifying the majority of infectious diseases; and, second, that ultimately effective control of these infections will follow widespread use of the time-honoured, sensible approach of environmental sanitation.

Each of these assertions is open to debate. It is true that building the Aswan high dam in Egypt increased the already high risk of developing schistosomiasis in Egypt, but that has to be balanced against the manifest advantages of building such a dam. Whenever such enterprises are contemplated, the environmental health implications should be considered by international experts convened by the World Health Organization and the Food and Agriculture Organization. Determined efforts should be made to prevent such health problems; this would be a useful way to use international aid funds, which are often squandered.

Does control of disease necessarily require the time-honoured, sensible approach of environmental sanitation? Nobody would argue against sanitation. But with labour and fossil fuel costs escalating, there is no realistic prospect of eliminating malaria from Africa by ditches and drains or tsetse flies by scrub clearance. Despite 30 years of the Chinese "People's War against the Snail" and Chairman Mao's poem "Farewell to the God of Plague", schistosomiasis still exists in China.

It is more realistic to employ whatever

weapons are at hand. Environmental sanitation should be used wherever it is feasible, and coupled with health education. Immunization too, has its place. Smallpox was eliminated by a vaccination campaign carried out with military determination, and immunization markedly reduced the incidence of poliomyelitis in industrialized countries. Chemoprophylaxis and chemotherapy are also useful. They greatly reduced the prevalence of yaws and have protected millions of people and domestic animals from parasites. As in the case of chemoprophylaxis against filariasis in

Samoa, quoted by Desowitz, knowledge of the social organization of the people helped.

The essays of Desowitz provoke these and many other thoughts. They are recommended to students of medicine, public health and biology, to administrators of international health programmes and to those who enjoy a good read about science and the world in which we live. □

Tony Allison is Director of the Institute of Biological Sciences at Syntex Research, Palo Alto.

greatest of all archaeological excavators General Pitt Rivers". It was to those largely forgotten publications from the end of the last century that Wheeler turned, and with their aid developed systems of recording, with emphasis upon the interpretation and drawing of stratigraphic sections, which transformed excavation in Britain from a pastime into a discipline.

Wheeler's most personal contribution was an unfailing grasp of what he called the Tactics and Strategy of excavation. Those same qualities which produced a first-class soldier (he rose to the rank of Brigadier in the Second World War and fought at El Alamein) and highly effective administrator (Director General of the Archaeological Survey of India) made him a masterly field worker, with an unrivalled perception of the real objectives of the work in hand and of the most appropriate measures for their achievement. Indeed, it could be argued that with Wheeler excavation became to archaeology what experiment is to the physical sciences — the opportunity through clear thinking and careful planning to try out new ideas and to test old ones. Of course British archaeology overseas had other pioneers of field method besides Wheeler — Sir Flinders Petrie, as Wheeler himself acknowledged, or Sir Leonard Woolley or Sir Max Mallowan. But Wheeler was the unflagging advocate of sound field method as the essential basis of further work, and his teaching in the years after the First World War promoted in Britain an awareness of stratigraphic principles which persists today and is ultimately sounder and more productive than the metrical thoroughness of the German school or the statistical enthusiasm of the American.

Jacquetta Hawkes does not perhaps evaluate as highly as she might this particular achievement, although she describes Wheeler's individual excavations fully and well. It deserves to be set in a wider context. For it would be quite possible to see Wheeler's keen sense of problem in fieldwork as anticipating in some respects the deliberately problem-orientated approaches of the "New Archaeology", which grew up in the last decade of Wheeler's life. It is not the inclination of either approach to set great store on facts just for their own sake. Both see the aim of fieldwork as the verification or testing of hypotheses through the gathering of fresh material, and recognize that it is ideas and problems which should determine the excavation strategy. This clear sense of priorities makes *Archaeology from the Earth* an inspiring introduction to fieldwork; still the best introduction, in my view, although many new techniques have been introduced over the past 30 years. And while the author rightly stresses Wheeler's ability to imagine and bring to life the people behind the archaeological record, his encouragement of the technical specialisms of archaeological science — whether conservation or radiocarbon

The making of an archaeologist

Colin Renfrew

Mortimer Wheeler: Adventurer in Archaeology. By Jacquetta Hawkes. Pp.416. ISBN 0-297-78056-5. (Weidenfeld & Nicolson: 1982.) £10.95.

"TURNING out of Pall Mall, I was transfixed by the steely gaze of Mr. Augustus John. 'Hullo, Rikki,' he said; 'still digging?'. 'Hullo, Augustus,' I replied; 'still sketching?'. With these words Sir Mortimer Wheeler (Rik to his friends) began *Still Digging*, his vivacious autobiography, published by Michael Joseph in 1955. Its subtitle, "Adventures in Archaeology", is recalled in that of the biography of Jacquetta Hawkes, now published six years after his death.

There is no more equivocal gift to the biographer than a really good autobiography. The freshness and energy of *Still Digging* make it lively reading today, and its very excellence offers Jacquetta Hawkes something of a challenge, to which she has risen conscientiously and above all sympathetically, making considerable effort to get behind the public persona and reveal the human being.

It is indeed for that persona that Wheeler is still most widely remembered. For those too young to have been viewers, it is difficult to convey the national impact of his charismatic role, back in the 1950s, in that remarkable television programme *Animal, Vegetable, Mineral*. Both Wheeler and the question master Glyn Daniel became household names — they were elected "TV Personality of the Year" in 1954 and 1955 respectively. This was *haute vulgarisation*, in Glyn Daniel's term, popularization at its best and with a serious purpose. It made archaeology both better known and more widely understood, laying the foundation

in Britain not only for more substantial subsequent television coverage (the *Buried Treasure* series, and then *Chronicle*) but also, it could be argued, for the rapid growth over the past 20 years of archaeology as a degree subject which is numerically strong in a number of universities. As the author stresses here, Wheeler was, in his early excavating days in the 1920s, a pioneer of what today might be called public relations. But of course it was very much more than this: Wheeler saw his discoveries in human, personal terms and he never lost the gift of catching the imagination of the non-specialist. He saw, too, that if the ultimate aim of archaeology is to inform us about the human past, it is the duty of the academic not only to research but to communicate. As he wrote in his best book, *Archaeology from the Earth* (Oxford University Press, 1954): "In a simple, direct sense, archaeology is a science that must be lived, must be 'seasoned with humanity'. Dead archaeology is the driest dust that blows".

In his later years, as Secretary of the British Academy, Wheeler made substantial contributions to the development of archaeology in Britain and indeed to the standing of the humanities in general. The British Academy, a much more recent creation than the Royal Society, has even now not quite achieved the active leadership within its own field which its elder sister has among the sciences. That the two are nonetheless at least comparable owes much to Wheeler's leadership during his tenure as Secretary from 1949 to 1968, as Jacquetta Hawkes describes most effectively.

Undoubtedly, however, Wheeler's main claim to lasting fame rests with his work as a pioneer of archaeological field methods. His incisive and systematic approach to the problems of excavation was not entirely new — as he generously acknowledged in the preface to *Archaeology from the Earth*, many of the methods and principles which he used were derived "from those of the

● The third edition of *Digging up Bones* by D.R. Brothwell, published by the British Museum (Natural History) and Oxford University Press, appeared earlier this year. Price in paperback is £8.95.

dating — and his recognition of their place within a more disciplined approach are at least as important.

The author's emphasis is perhaps a more personal one, although she copes well and expertly (being an archaeologist herself) with Wheeler's excavations and research. She is concerned to see Wheeler as a person, dealing as much with the private life (and many loves) as with the public achievement. She begins: "Mortimer Wheeler will rise from these pages as a Hero

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Mortimer Wheeler — creator of a discipline.

figure. Of that I feel sure even as I write the very first words of my book". Certainly the remarkably varied career — yet single-minded in its devotion to archaeology — is admirably told. Her earlier chapters and description of the war years inevitably owe a great deal to the autobiography. But surprisingly, perhaps, it is the portrait of Rik Wheeler in his old age, as he is remembered by many today (including myself) which seems the most compelling, indeed moving. Those same qualities of panache, of awareness of the effect of the moment, which may have raised eyebrows (and indeed aroused some enmities) earlier, and which caught the attention of millions of viewers in the 1950s, made him the most splendidly vivacious good company in his seventies and eighties, and a friend and encouragement to many half a century younger than he. They come over well in this affectionate portrait.

Wheeler came to archaeology with a military briskness of mind which helped him to make the subject, in the scientific as well as the organizational meaning of the term, a discipline. Perhaps just as important, he brought with him a zest and a gusto which made it, in a real and un-trivial sense, entertaining. □

Colin Renfrew is Disney Professor of Archaeology at the University of Cambridge.

Grounds for doubting the pessimists

P. D. Henderson

The Ultimate Resource. By Julian L. Simon. Pp.412. US ISBN 0-691-09389-X; UK ISBN 0-85520-440-0. (Princeton University Press/Martin Robertson:1981.) \$14.50, £9.50.

WRITING in 1933, Keynes said of Malthus's *Essay on the Principle of Population* that "it attracted immediate attention, and the warfare of pamphlets instantly commenced . . . which for 135 years has never ceased". Hostilities have continued during the half-century since Keynes wrote, and Julian Simon's new 400-page volume is a reminder that not only pamphlets are involved. The issues that Malthus raised so provocatively in 1798 are still live, unsettled and highly contentious.

The main issues are two: first, the extent to which it is possible for human societies to achieve material progress; and second, the effects of population changes on their prospects of doing so. In relation to these, it is still possible to think of the main battleline as drawn between Malthusians and anti-Malthusians, though as in other spheres of conflict there are wide differences of view within each of the two factions.

Following in the steps of the master, modern Malthusians are pessimistic on both issues, or at least inclined to emphasize dangers and limitations rather than opportunities. As to material progress, the possibilities for growth in output are seen as restricted in any given society, and for the human race as a whole, because of constraints on the availability of land or natural resources. From this arises the threat of overpopulation. It is admitted that population growth will create a larger labour force — more hands, as well as more mouths — and thus tend to raise the level of potential output; but given the constraints imposed by nature, not much can be expected from this positive effect. Thus higher rates of population growth will often be a drag on material progress. Conversely, lower population growth may hold the key to sustainable prosperity.

By contrast, anti-Malthusians take a more sanguine view. They are sceptical about limits to growth, while at the same time many of them stress the positive effects of population increase on a society's capacity to produce goods and services. These are the two main themes of Simon's interesting and readable book. The author pours scorn on current pessimism and prophecies of doom — despite one or two favourable references to Malthus himself — and the book may be described as a scholarly polemic.

Most of the argument is directed to the two issues identified above. Part One ("Towards Our Beautiful Resource Future") considers the possible scarcity of natural resources, and deals successively with minerals, food supplies, land, energy,

environmental pollution and conservation. On all of these Simon points out, with some well chosen and carefully presented statistics, that the evidence of recent decades is more encouraging than is generally realized. Thus minerals, food products and energy supplies have over the past century become less costly in relation to wages; world food supplies have grown more rapidly than population; while the expectation of life, which he regards as the best single indicator of pollution, has risen all over the world. In Simon's judgement, all these favourable trends are likely to be maintained into the indefinite future: in the case of the cost of raw materials, he makes an engaging offer to back his assessment with money, by accepting bets from readers who take a different view. He thus rejects completely the notion that economic growth is constrained by the availability or cost of resources.

Part Two is mainly concerned with the effects of population growth on average income per head. In most of the economic-demographic models that have been developed over the past 20 to 30 years, these effects are unfavourable, largely because it is assumed that with more children in the average family the rate of saving and capital accumulation will be lower, and that a higher proportion of capital investment will go towards duplication of facilities (more schools, for example), rather than increasing capital per head. As against this, Simon makes three main points. First, larger families may increase the willingness of individuals to save and invest. Second, there are likely to be significant gains from economies of scale, since total output will be higher. Third and most important, output per head will benefit from "the contribution of additional people to our stock of useful knowledge" (p.196): a larger population means not only more mouths and hands, but also more brains. His broad conclusion, which he argues is consistent with the facts of economic history, is that moderate as distinct from zero rates of population growth are likely to have favourable effects on income per head, at any rate in the longer run, in rich and poor countries alike.

In Part Three ("Beyond the Data") the argument is taken further than the issue of material progress alone, and other values are brought in. Simon holds that enabling an extra person to live and enjoy life has value in itself, and that the well-being of a society has to be defined with reference not just to the average standard of living of its members, but to their total numbers also. He makes a sharp attack on the values and assumptions of anti-natalist bodies, such as Planned Parenthood, and argues that there is no justification for trying to reduce rates of population growth, as distinct

from enabling parents to obtain the size of families that they would wish.

Throughout the book Simon's own position, as summarized just above, is contrasted with the latter-day Malthusians whom he wishes to discredit. There are numerous quotations, some from learned articles and monographs, but most from popular books, newspaper articles and advertisements, designed to illustrate the whole range of contemporary pessimism. The style is kept deliberately informal and colloquial, while the author's convictions are often set out in a highly personal way. Now and then this mode of presentation becomes a little trying, but in general the argument is clearly and effectively conveyed, without either condescension or oversimplification. In view of the range and complexity of the topics that are covered, this is a considerable feat.

In judging the merits of Simon's argument, two distinct tests are relevant, corresponding to the two aims which the book tries simultaneously to achieve. The first aim is destructive: to undermine and cast doubt on current neo-Malthusian pessimism. The second is to offer an alternative view of the world, an acceptable analysis of events, relationships and trends. Not surprisingly perhaps, the author succeeds more completely with the former task than with the latter.

The process of demolition is carried out with skill and resource. In part, the prophets of doom are discredited by the various excerpts from their own works, many of which make sorry reading. But the main weight of the argument rests on the evidence of the past and on general economic reasoning, both of which are often neglected or undervalued by the pessimists, and which the author presents extremely well. Unfortunately, the views that he attacks are widely held and influential. By challenging them so effectively Simon has performed a real service.

When it comes to his alternative view of the world, there is more room for doubt. At a number of points the argument seems to me to be too unqualified or incautious, while in some respects it is less complete than it should be. For example, his generally admirable discussion of natural resources is marred by a failure to consider the possibility that oil is a special case. Because of what are often presented as facts of geology, it may be that the period of cheap oil was a historical accident, a bonanza which cannot be repeated. If this is so, and given the unique convenience and properties of petroleum as a fuel, a secular rise in the real price of oil is not at all to be ruled out; and it is insufficient to dismiss the increases of the 1970s, as Simon does, as being due simply to the machinations of OPEC. Moreover, because of its role in the world economy, and the uneven distribution of productive capacity and reserves, the effects of oil price increases are potentially much more damaging than

in the case of other raw materials. Simon does not go into this, and his two chapters on energy are superficial.

As to population, a clearer distinction could have been made between two aspects of the question: the absolute size of a population, and its rate of growth. More space should also have been given to two arguments against rapid population growth: one is the possible congestion arising from higher densities, while the other is that a rapidly growing labour force may lead to a more unequal distribution of income. Again, Simon states correctly that a judgement on the net advantages of population growth may depend on how one weighs more distant costs and benefits in relation to those which are nearer in time. But he says nothing about how this valuation of time should be made, nor about the suitability of using for this purpose the actual rates of interest that are established in financial markets.

More worrying is an occasional tendency to rather serious overstatement. Thus in the concluding chapter, we are told that "The standard of living has risen along with the size of the world's population

since the beginning of recorded time" (p.345). This is false. One can find phases of history — as for example in England from the mid-thirteenth century till around the end of the seventeenth — when a rise in population is thought to have led directly to a fall in real wages, and vice versa: the simple Malthusian model seems to fit the facts. It is probably only in the modern era that this inverse connection has been broken, first in the industrial countries and more recently in other parts of the world: possibly it holds good even now in some of the poorest and least developed countries. Here and at some other points of the argument, Simon has yielded to temptation: he makes overconfident generalizations about relationships that are subject to considerable variety and change.

Despite these reservations, the book is much to be recommended. It is possibly the best available modern treatment of the issues that Malthus raised. □

P. D. Henderson is a Professor in the Department of Political Economy at University College London.

The hunting of the rorqual

Arthur Bourne

The History of Modern Whaling. By J.N. Tønnessen and A.O. Johnsen. Pp.798. UK ISBN 0-905838-23-8; US ISBN 0-520-03973-4. (C. Hurst, London/University of California Press: 1982.) £22.50, \$45.

THE POPULAR image of whaling as the harpooner, lance in hand, standing bravely on the prow of a row-boat ready to do battle with Leviathan, is enshrined in the writings of authors like Melville and Bullen. The picture is a romantic one but, as both those writers knew, whaling was far from romantic. It was — and its rump is — a tough business, firmly based on hard commercial interests. If there is little romance in whaling, as compensation there is a world of fascination and of no period is this more true than when the new whaling industry was emerging from the old.

The modern epoch really began in February 1864 when a small steam-powered vessel, the *Spes et Fides*, sailed out of Tønsberg in south-eastern Norway for the whaling grounds off Finnmark. It is the story of the unfolding of that modern history which is the subject of *Den Moderne Hvalfangsts Historie*, of which the volume under review is a much-condensed English translation. In it the authors have cleverly drawn together the strands of the technical, sociological and financial developments of that period.

The birth of the modern whaling industry will always be associated with one man, the owner and skipper of the *Spes et*

Fides, Svend Foyn. Why he was to be the leading force and why Norway should have been the birthplace are interesting questions. As the authors show, the Americans dominated the old whaling scene and were experimenting with new methods at the same time as Foyn. Even his idea of combining an explosive harpoon fired from a cannon mounted on a steam-powered boat was not original, and he was not the first man to realize that the future of whaling lay in the chasing, catching and processing of the great rorquals. The real answers to those questions lie not only in the technical inventiveness of the various contenders — of which there were many — but also in the social and economic changes that were occurring at the time.

In the middle of the nineteenth century the Americans were looking increasingly inwards. The opening of the West with its vast prairies and potentially enormous mineral wealth seduced capital and labour away from the coast and the sea. In Norway the situation was very different. With little hinterland and an increasing population which the nation's industry and agriculture could not sustain, thousands of Norwegians, in common with many other Europeans, sought their futures in North America. By one of those quirks of fate, America's retreat from the sea opened up an opportunity for the Norwegians left at home. They were quick to fill the gap, and, combined with the expansion of sea-trade generally, this provided the stimulus for the

growth in Norway's maritime industry.

At the same time the economics of whaling were propelling technical developments along parallel lines in several countries. The stocks of right whales — "right" because they were slow swimmers and floated when killed — had been so decimated by the middle of the century that they were unable to support a viable industry. The interest of the whalers turned to the rorquals which "teemed in the seas", but which were too swift for the row-boats and, moreover, sank when killed. One of the first to experiment with explosive harpoons was the American T.W. Roys, who invented a harpoon-throwing gun and set out to see "if a cargo of oil could be obtained from whales previously unavailable to mankind". There were other innovators and other developments, one of which was, the authors believe, of decisive importance to Foyn's own ideas. This was the grenade harpoon invented by George Welch in 1867, three years before Foyn's final patent was granted. In many countries similar inventiveness was displayed, but Foyn was the most persistent. He believed in the potential of the rorqual fishery — even during his experiments with the *Spes et Fides* when failures lost him money, he kept going. He was in the end the right man at the right time.

The first of the rorqual stocks to follow the right whales to extinction was, not surprisingly, the Finnmark stock. From then on it was for the whaling companies a matter of geographical expansion and greater sophistication in techniques and ship design. The whalers extended their activities to wherever rorquals could be found and eventually to the waters surrounding the Antarctic continent. It was there that the greatest harvests were to be culled, and where the technical ingenuity of the whalers was again stretched.

The problems were to stay at sea long enough to justify the cost of fitting out a fleet, and to build more powerful and faster catchers. The floating whale-factory, heralding the pelagic whaling industry, extended the range of the catchers, and the introduction of off-loading the oil into tankers enabled the fleets to remain at sea for months. From 1937 and the launching of the Japanese-built *Seki Maru*, the catchers became progressively more powerful. However, no matter how long the whalers were on the oceans, nor how sophisticated their equipment, if the whales were not there they could not be caught. Inevitably the Finnmark story was to be repeated again and again, and nation after nation, including Norway, had to abandon pelagic whaling — political ingenuity in conserving whale stocks fell far short of the technical expertise in destroying them.

The whaling industry is neglected by the student of the development of the industrial nations and yet it reflects the social and economic conditions of the

times most vividly. It also provides clear insight into the way those conditions forced men to seek means of utilizing resources that had hitherto been out of reach. The story of the whaling industry shows too how difficult it is to maintain an industry based on natural resources, even renewable resources, if immediate commercial considerations over-ride the long-term benefits of rational exploitation and investment.

The authors of this book have achieved a remarkable piece of condensation and R.I.

Christophersen must be congratulated on a fine translation. The only sad thing about this volume is that it is unlikely that we shall ever see an English-language version of the original four-volume work — and that it is equally unlikely that mankind will learn from the mistakes it so amply documents. □

Arthur Bourne was an observer at the International Whaling Commission from 1964 to 1968. His great-grandfather was a contemporary and competitor of Svend Foyn.

Cartography: from astrolabe to orbiter

Helen Wallis

The Mapmakers. By John Noble Wilford. Pp.414. UK ISBN 0-86245-041-1; US ISBN 0-394-46194-0. (Junction Books/Knopf: 1982.) £9.95, \$20.

A NOTABLE theme in the history of human endeavour is the record of man's survey and mapping of the Earth, the Moon and (by spacecraft) of the nearer planets. Even the earliest surviving records show a surprising degree of sophistication in the cartographic arts. Thus an example of early cadastral mapping has been identified in the depiction of a village found in rock carvings of the Valcamonica in northern Italy and dating from the second and first millennia BC. The Babylonian world map of about 500 BC is one of the earliest depictions of the cosmos. And two silk maps recently recovered from the Ming tombs show the advanced level reached by the Chinese in the second century BC.

With the development of the geographical sciences in classical Greece, the coordinate system for the mapping of the world was established which is still the accepted method today. Even in the so-called "dark ages" of Europe, a continuing tradition can be traced from the surveys of the late Roman empire through mediaeval times.

The European Renaissance brought about a revolution in mapmaking. The newly invented techniques of engraving

made possible the "exactly repeatable picture", and in the late fifteenth century led to the establishment of the mapmaker's craft and the map publishing industry. At the same time European ventures of exploration set in motion the mapping and surveying of the world. A comparable revolution has taken place in the past 50 years with the development of aerial surveys, automated cartography, and mapping from and in space.

The history of the complicated processes involved in these developments can be told in various ways. The role of the mapmakers themselves is often subordinated to the record of the artefacts, the atlases, maps and globes. John Noble Wilford in his new history of cartography sets out to correct this bias. In *The Mapmakers* he tells the story of "the multitude of diverse characters", ranging from scholars and scientists to clockmakers, schoolteachers and spies, aviators and technicians, who stand out in the history of cartography. Although he has not spent a lifetime working with maps as such, his professional post as science correspondent for *The New York Times* makes him well qualified to identify the pioneers and innovators in the history and explains his concern to bring out the drama of personal triumph. As professional mapmakers normally write their reports in "the bloodless language of the specialized

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Surveyors at work in the eighteenth century — detail from a map of Warwickshire by Henry Beighton, 1728.

journals'', he supplements written sources for modern surveying with interviews with the practitioners themselves, at NASA, the US National Geodetic Survey and so on.

No single volume on the history of cartography can be comprehensive. It is important therefore to note that Wilford's book is really a history of the attempts to measure, survey and map the Earth, the Moon and the planets. The definition of cartography implied in this treatment is that of the United Nations, therefore, and not that of the International Cartographic Association (which excludes survey from its terms of reference). What is particularly valuable is the integration of the historical and the modern. Most histories of mapmaking end in 1900 or earlier, and leave to others the recording of recent developments. Thus the first two parts of the book deal with the broad trends and signal achievements of cartography and geodesy prior to the twentieth century. The final two parts include some earlier history but are focused on the twentieth century.

Another, less welcome, aspect of the treatment is an evident "Western" bias. Like many other histories, the book is "ethno-centred" (to quote the current term), dealing largely with the achievements of Mediterranean and European peoples. Arab geographers such as Idrisi, for example, are conspicuous by their absence, and the remarkable achievements of Chinese cartographers appear only in the pages dealing with the maps from the Ming tombs. In so focusing the story Wilford follows Lloyd A. Brown, whose *Story of Maps* (Little, Brown; 1949) provided the classic account of mapmaking to 1800. The similarity in treatment and in the selection of illustrations for the early period explains why some parts of Wilford's book (for example, the sections on the Cassini and John Harrison) will seem almost familiar to those well acquainted with Brown. Wilford's detailed account of the survey and mapping of India and North America extends the compass, however, in addition to the fact that the time span covered by Wilford is much greater.

The book is illustrated by reproductions of some of the most important maps and by photographs, but documentation of the illustrations is inadequate. The Babylonian world map (p.11), for example, is in the British Museum, not the New York Public Library (p.415), as the reader might suppose from the credits. Sources for information in the text are, for the most part, not cited, but bibliographical notes are provided for each chapter.

In general, however, Wilford's *Mapmakers* can be warmly recommended as an essential addition to library reference shelves, and as an excellent introduction to mapmaking past and present for geographers and map lovers alike. □

Helen Wallis is Map Librarian at The British Library, London.

Ancient woods in modern times

John Andrews

Woodland Conservation and Management. By G.F. Peterken. Pp.328. ISBN 0-412-12820-9. (Chapman & Hall: 1981.) £25, \$49.95. *Ancient Woodland: Its History, Vegetation and Uses in England.* By Oliver Rackham. Pp.402. ISBN 0-7131-2723-6. (Edward Arnold/University Park Press: 1981.) £50, \$99.95.

DURING the past 30 years, more ancient mature woodland has been destroyed in Britain than in the preceding four centuries. The direct cause of this has been the parsimonious attitude of post-War government towards forestry and a corresponding benevolence to agriculture, with an increasingly phoney system of cost-benefit appraisal favouring over-production from farming while squeezing forestry into short-rotation, softwood monocultures whose main product is pulp.

It is a tribute to the Forestry Commission's technical skill that it has succeeded in developing techniques of wood production under uniquely difficult conditions in the uplands, but no credit to foresters in general that they have fostered the destruction of Britain's remaining native woodlands, both by encouraging their replacement by conifers and by meekly acceding to their conversion to farmland. Perhaps conservationists too must share the blame, for failing to protest sooner and louder: not least the Nature Conservancy Council which should lead the movement, and yet still shies away from producing a policy on forestry after nearly a decade of indecision.

Fortunately, many NCC staff are a good deal harder-nosed than their politically-appointed Council. *Woodland Conservation and Management* is written by George Peterken, a member of NCC's scientific team, and ventures some provocative views. Thus,

private landowners, who enjoy a disproportionately large share of national wealth, should be prepared to bear a disproportionately large share of the sacrifices needed to provide national needs for amenity and nature conservation. Indeed, the continued political acceptability of private landowning may depend on recognising these and other public responsibilities

— a view gaining currency among conservationists patronizingly advised to leave protection of the countryside to the good will of the landowning community.

The book, however, is primarily a scientific work. It commences by examining the origins, management and ecological characteristics of British woodland, demonstrating the special value of ancient, primary woodland to a wide range of species which have persisted in such sites — and nowhere else — despite many centuries, perhaps in some cases several millennia, of human management.

The second part of the work considers the types of semi-natural woodland, their classification and management. Woodland classification in Britain is still bedevilled by the lack of a generally accepted system and here the effort is made to present one for use by any competent naturalist or forester. Such a classification has an important practical value, providing the basis for identification of woodland sites of similar type so that comparative conservation and management judgements can be made.

About a third of the book is devoted specifically to woodland nature conservation. There is much helpful advice on the recording and assessment of woodlands and on management methods, both where conservation takes priority and where the first objective is production. Though the emphasis is on native woodlands, the pros, cons and opportunities presented by the new conifer forests in the uplands are not neglected.

The book is well and economically written. Though some sections are technically complex, it contains much of general interest, not least a clear and concise statement of the objectives and priorities of nature conservation — matters on which many foresters and some naturalists seem confused or misinformed. However, as an ornithologist, I must regret the author's heavy botanical bias: though birds have great powers of dispersal, many survive best in ancient woodland; and the absence of any reference to Britain's only endemic bird — the Scottish crossbill, which appears to be largely confined to Caledonian pine forest — cannot be lightly forgiven.

In *Ancient Woodland*, Oliver Rackham approaches the subject from a different standpoint. In the author's words, the main theme is

the self-renewal typical of English woodland history: the quiet, usually unremarked regrowth of the young shoots, automatically replacing the woods every time they were cut down, decade after decade and century after century.

But it is more than that. Dr Rackham's special skill lies in unravelling man's relationship with woodlands and, in scholarly yet entertaining style, he deduces from them aspects of rural social history since pre-Roman times.

Though the book contains examples from many parts of England and the Continent, it is based largely on the author's own studies in East Anglia. This region is a splendid study area, containing an unusual variety of woods which have been well documented — paradoxically the outcome of early clearance which meant that the remainder assumed special importance for local use and was systematically conserved until recent years.

Roughly the first quarter of the book covers woodland classification, soils, plant and tree communities; in some measure this overlaps with Peterken but the emphasis on man's role in shaping the woods gives a usefully different slant. Evidence from pollen analysis, soil structure, earthworks and archaeological remains, vegetation and tree rings, oral tradition, paintings and written records is brought together to present woodlands not only as places of beauty, rich in wildlife, but also as a part of our heritage — as much a reflection of past human skills, aspirations and shifts of power as any cathedral.

Moving on to the human uses of woods, the author leads us from Neolithic times through the extensive records of Domesday to

the mad world of the 1970s, in which the price of ordinary oak-trees is as low, in relation to the

value of money, as at any time since the fifteenth century; the price of oak timber is higher, again in relative terms, than at any time in history; and prices paid for underwood, provided the seller knows where to find the buyer, seem to be roughly equal to the previous all-time maximum in the eighteenth century.

Nearly half of the text deals with individual tree species, examining their distribution, prehistory and history, modern ecology and conservation. With many references to specific sites, one is presented with the opportunity to go and see for oneself — a temptation I will certainly find irresistible when the sun shines, just as dipping into the book will be a continuing pleasure on rainy days. □

John Andrews is Head of Conservation Planning at the Royal Society for the Protection of Birds, Sandy, Bedfordshire.

The malaise of modern medicine?

Robin Weiss

The Diseases of Civilisation. By Brian Inglis. Pp.371. ISBN 0-340-21717-0. (Hodder & Stoughton:1981.) £10.95.

BRIAN Inglis is a respected journalist and author who has written several informative books on the history of medical practice, orthodox, homeopathic and supernatural. His latest, *The Diseases of Civilisation*, is subtitled "An indictment of traditional methods of medical treatment and a plea for a completely new approach". Inglis surveys the major diseases affecting the Western world today — heart disease, cancer, mental illness, common infections, "auto-immune" and degenerative diseases — and concludes that modern medical science and practice has failed to live up to the high hopes of the immediate post-War period. Rather than question how realistic these expectations ever were, generated by journalists like himself as much as by the medical profession, Inglis concludes that the medical establishment is to blame for the intractability of ill-health in modern times. He depicts the medical profession as a self-perpetuating oligarchy composed of too many narrow specialities which dominates medical education and research, distorting both to serve its own ends and thereby obstructing public health. The food, tobacco and pharmaceutical industries are the accomplices in this medical crime, abetted by Governments that shy away from confrontation with an entrenched medico-industrial complex. Inglis argues that this unholy alliance has prevented the medical profession from embracing and acting upon discoveries of the important roles of diet, stress and other social factors in the causation of disease, so that resources which should be channelled

into primary health care are deflected into the ever-increasing expense of high technology hospital treatment.

Much of this argument is regrettably only too true. The medical establishment is extremely conservative and profoundly suspicious of social or "unorthodox" approaches to disease. As Thomas McKeown has long argued, medical education is almost entirely geared to the treatment of acute illness as presented in teaching hospitals; the underlying causes of disease, other than specific pathogens, receive scant attention and the very concept of health is hardly raised during the training of doctors. Governments faced with butter mountains and tobacco revenues usually duck environmental and health issues when confronted by powerful lobbies — witness President Carter dismissing Mr Califano from the Department of Health, Education and Welfare, and Mrs Thatcher reshuffling Sir George Young out of the Department of Health and Social Security as soon as they tried to take significant action on curbing the promotion of cigarette consumption.

Brian Inglis has tackled a most important subject — the limitations and misconceptions in the modern Western approach to illness and health — yet I find his book disappointing and his message pernicious. Posing as a carefully researched documentary, it is more a polemical tract in the tradition of Ivan Illich's *Medical Nemesis* in which Inglis overstates his case to the point of being anti-science. His main thesis is that psychosocial stress is not taken into account in the investigation, management and prevention of diseases of civilization. He cites Stewart Wolf's important finding

that heart disease hardly occurred in Roseto, Pennsylvania, while the inhabitants maintained a cohesive, conservative society, although they ate a typical Western diet. Wolf himself in *Social Environment and Health* (University of Washington Press, 1981) views disease as maladaptation to changing social conditions. Inglis concludes that it is therefore wrong to treat disease by orthodox methods. He proceeds to select the failures of modern health care and to ignore all the successes. He would have us believe that all heart surgery is a wasted resource, that no cancer patients benefit from chemotherapy, that mental hospitals should be abolished and that immunization plays little role in preventing infectious diseases. He states that "the net value of orthodox medical treatment is far less than it was twenty-five years ago, and is declining".

The chapters I find most objectionable, no doubt because they deal with my own subjects of study, are on infectious diseases and on cancer. Inglis dismisses research into the identification and control of infectious agents in disease (for example in Legionnaires' disease and in malaria) as "mechanistic heresy" obscuring psychosocial elements. He also shows a remarkable prejudice against virologists, imagining that they narrowly pursue infectious organisms without reference to epidemiology and pathogenesis. Elsewhere he supposes that virologists have sought to monopolize cancer research, preventing immunologists investigating "cancer immunotherapy", whereas it is the virus-induced cancers that may yield first to immunological means of prevention or therapy. The eradication of smallpox, the resounding medical triumph of the 1970s, is not cited at all in Inglis's book, yet its success depended on the very combination of preventive vaccination and thorough epidemiological investigation that orthodox medicine provided where 2,000 years of ayurvedic methods in India failed. The successful outcome of virological research leading to immunization against poliomyelitis earns a grudging mention, with the extraordinary suggestion that the rare but devastating neurological damage caused by polio virus infection is more related to psychosocial stress than to the virulence of the virus strain. Polio and infectious mononucleosis are two particularly intriguing diseases of civilization in that children of the higher social classes are the ones most at risk; the epidemiological factors underlying this phenomenon are not probed by Inglis, presumably because they do not fit his view of how virologists conduct research.

Inglis's criticisms of orthodox cancer treatment resemble those which have recently occupied the columns of the *Washington Post* and the diatribe by Dick Richards (*The Topic of Cancer: When the Killing Has to Stop*; Pergamon, 1982) against those who "cut, burn and poison"

cancer patients. Few would deny that patients sometimes receive excessive or inappropriate treatment; but having berated specialities, Inglis does not see that this is most likely to happen when cancer patients are treated by non-specialists. Inglis blames the medical establishment for the lack of progress in treating cancers of the lung and breast. Yet he makes no reference whatsoever to the successful treatment and cure of many paediatric cancers and leukaemia, Hodgkin's disease, choriocarcinoma or testicular cancer, which all result from the development of the aggressive forms of therapy which he so derides. True, these are rare types of cancer in terms of total incidence of malignant disease; but they are major causes of serious illness in children and young adults, with a high social and personal toll. There are common cancers too — of the skin, for example — for which diagnosis at an early stage undoubtedly reduces mortality very significantly. Inglis also entirely ignores palliative treatment, assuming that surgery, radiotherapy, chemotherapy or endocrine treatment invariably cause more suffering than leaving the cancer patient untreated by orthodox methods.

Inglis's account of cancer research funding policy is also seriously misinformed. He states that because the British cancer research charities "are largely controlled by specialists, naturally they favour projects which help them in their hospital work, providing them with better-equipped operating theatres, improved radiotherapy machines, more powerful drugs". This is nonsense, as no research project is more difficult to obtain funding for than one that smacks of a medical service role. Following the *British Medical Journal's* peculiar outburst two years ago, Inglis reiterates that less than 2 per cent of the British charities' expenditure is directed towards research into cancer epidemiology and prevention, an absurd underestimate even if one discounts the potential of basic research for future prevention. He thinks the public should be given more opportunity to influence the way in which cancer research funds are spent. This is a laudable sentiment, though difficult to implement;

moreover, "the public" will almost always choose to fund high-technology hardware, such as CT scanners, in preference to research into the psychosocial aspects of cancer that Inglis rightly wishes to promote.

In criticizing the medical management of the mental patient Inglis is on stronger ground, in that no single orthodoxy has emerged. From his account one does have the feeling that the theory and practice of treatment for mental illness has not advanced beyond the blood-letting era. Having heaped scorn upon the medical establishment's "obsession with diagnosis" for infectious diseases, Inglis here deplores the unreliable classification of mental disease, quoting large discrepancies between US and British psychiatrists. In fact, the WHO definition of schizophrenia can now be used to standardize diagnosis. As Julian Leff discusses in his recent monograph *Psychiatry Around the Globe: A Trans-cultural View* (for review see p.523), the incidence of schizophrenia is surprisingly uniform across different cultures, though its recognition as a disease state requiring treatment differs widely.

Inglis misses another opportunity for serious debate in his chapter on iatrogenic disorders. Quoting the Hippocratic dictum "At least, do no harm", he does not discuss whether the harm induced by treatment might be less than that of no treatment. Inglis justly reproaches the medical profession over drug abuse and carefree prescription, particularly in treating the symptoms of diseases of civilization without paying due attention to underlying causes. Yet in an earlier chapter he criticizes by-pass surgery for angina pectoris when chronic drug treatment might achieve the same result. One of the most notorious examples of iatrogenic disease (not quoted by Inglis) was the induction of leukaemia in patients treated with X-irradiation for ankylosing spondylitis. The radiotherapy undoubtedly caused enormous relief from the symptoms of a crippling, painful and eventually fatal hereditary disease; given the prospect of that relief, with the possibility of dying from iatrogenic leukaemia 15 years later, would Inglis have advised the spondylitis patient to spurn treatment?

Overall, Inglis has broached an important topic by questioning the basis of our current medical philosophy and its disregard for preventative health measures and social care. But I wish that he had not chosen to present orthodox medicine and research as the antithesis of genuine care and that he had researched his commentary more deeply. More importantly, I earnestly hope that any young man with testicular cancer or parent with a child suffering from leukaemia will not refuse orthodox treatment as a result of reading this book. □

Robin Weiss is Director of the Institute for Cancer Research, London.

Politics of pesticides

M.S. Swaminathan

Circle of Poison: Pesticides and People in a Hungry World. By David Weir and Mark Schapiro. Pp.99. ISBN 0-935028-09-9. (Institute for Food and Development Policy/Third World Publications: 1981.) \$3.95, £2.50.

Circle of Poison provides an insight into the process by which "someone in the underdeveloped countries is poisoned by pesticides every minute". The circle begins with the activities of leading multi-national corporations which dominate the seven-billion-dollar-a-year pesticide market. The links in the chain include loopholes in government regulations in the United States, the import policies of developing countries and the indifference of bilateral and multilateral donors, regional and international banks and UN agencies.

Weir and Schapiro argue their case cogently. For example they point out that at least 25 per cent of US pesticide exports are products that are banned, heavily restricted or have never been registered for use in the United States — "The Federal Insecticide, Fungicide and Rodenticide Act explicitly states that banned or unregistered pesticides are legal for export". Developing countries which receive hazardous pesticides recommend the application of such pesticides in a routine manner, partly due to ignorance and partly due to their being cheaper than less-toxic products. An instance has been cited where parathion is being used in Central America at a dosage 40 per cent greater than necessary. In contrast, through the adoption of better pest-management procedures, US farmers now use 35–50 per cent less insecticide than ten years ago, with no adverse effect on crop yield. An important consequence of the indiscriminate use of pesticides is the multiplication of pesticide-resistant insect species, leading to the use of still higher concentrations of the chemical, and so on.

At each stage in the process, the circle of poison has its victims — the staff in pesticide manufacturing plants, those who load and unload the chemicals, Third World peasants, workers and consumers, and finally everyone else in the world who eats food contaminated with pesticide residues. But undoubtedly the greatest impact is felt in the developing nations. According to some calculations, the rate of pesticide poisoning in such countries is more than 13 times that in the USA. The authors give examples from their investigations in Mexico, Central America, Pakistan, Indonesia and Papua New Guinea to illustrate how pesticides prohibited in the US reach the peasants of the Third World. Unfortunately, in these countries there are neither enough scientists to investigate the dangers arising from such pesticides nor is there a well-

Handlist for historians

The Science Museum, London, has recently published a vade-mecum for science historians, *Reference Books for the Historian of Science: A Handlist*. The bibliography has been compiled by S.A. Jayawardene with the intention of providing a volume to complement other bibliographic manuals, covering the primary and secondary sources of the history of science, by the inclusion of general reference books. The handlist includes some 1,000 titles, and is available from Library Publications, Science Museum Library, South Kensington, London SW7 5NH, price £2.50, £3 by post.

informed and well-equipped government machinery to prevent foreign pesticide makers from selling products which are banned in their home countries.

However the United States and other pesticide exporters also suffer the consequences of the dumping of banned chemicals in developing countries. According to the US Food and Drug Administration, about 10 per cent of the food items imported into the United States contain higher levels of pesticides than are permitted. Therefore, even on the basis of enlightened self-interest, it would be advisable for developed countries not to permit the export of chemicals which cannot be sold at home.

The authors also question the pathway of productivity improvement involved in "the green revolution strategy". By encouraging monoculture, losing genetic diversity and not promoting integrated pest-management schedules, the use of large quantities of pesticides has become essential. The authors suspect that the emphasis on this kind of high-yield technology is largely because of the interests of the multi-nationals — several of them are involved not only in the production and sale of pesticides, but also in the production and distribution of seeds and chemical fertilizers.

The circle of poison has become possible only because of a circle of collusion between national and international companies, banks and government and UN agencies, all playing a role in the spread of poisonous chemicals. Fortunately, in recent years, some steps have been made towards breaking the unholy triple alliance of pesticide manufacturers, importers and finance institutions. The authors propose several methods by which the fight against the proliferation of hazardous pesticides can be brought to a successful end. Action, they say, has to be taken both at national and international, and individual and institutional levels. Examples are cited of the kinds of effective checks being developed in the Philippines and Malaysia. The resolution of the United Nations General Assembly passed in December 1979, urging member states to exchange information on hazardous chemicals (and pharmaceuticals) that have been banned in their countries, and to discourage export of those products to other countries, is another move in the right direction. This has been followed by similar steps by OECD and other organizations. The Environmental Protection Agency in the USA is also becoming more vigilant.

David Weir and Mark Schapiro have rendered signal service by drawing attention to the deficiencies in existing regulations and practices in the trade of hazardous pesticides in both the exporting and importing countries. The book, therefore, deserves to be widely read by all connected with agricultural research and development as well as those involved in the development of public policies relating

to human health and environmental protection. Nonetheless, their account does have some flaws.

In the section entitled "More Food and Yet More Hunger", the authors try to place the blame for rural poverty mainly on developments in the improvement of agricultural production. This is not wholly correct since it is now widely accepted that the new technology involving the cultivation of management-responsive varieties is neutral to scale with regard to the size of the farm in which the technology is adopted. Hence, a small farm is not a handicap from the point of view of adoption of new technology. A small farmer, however, has many problems arising from the cost, risk and return structure of farming. The potential offered by a small farm for intensive agriculture can, therefore, be realized only by removing the constraints faced by small farmers. The action required is in the realm of public policy and not technology. Technology can only help to maximize

output from the resource endowments of each country; poverty can be minimized only through appropriate political decisions in the fields of land and livestock reform, asset transfer and social security. Unfortunately, terms such as "miracle seeds" have created the impression that new technologies can provide miraculous solutions to problems of poverty, social inequality and injustice. It would have been prudent to have avoided the temptation of making a few casual and over-simplified statements on rich-poor equations, since such problems deserve in-depth analysis. This would have lent greater authority and credibility to the principal message of the book — that the temptation to make profit out of poverty and ignorance can be countered only through widespread awareness of the techniques used for this purpose. □

M.S. Swaminathan is Director General of the International Rice Research Institute, Manila, Philippines.

Whence the power of the medicine-man?

Robert Ubell

The Clay Pedestal: A Re-examination of the Doctor-Patient Relationship. By Thomas Preston. Pp.226. ISBN 0-914842-68-4. (Madrona: 1981.) \$12.95.

THOMAS Preston, the author of this book on modern medical practice in the United States, is a cardiologist with his heart in the right place. Acutely aware of the way doctors play musical chairs with their patients, he rises to their defence, sympathizing with the patients, castigating the physicians.

Preston's analysis closely follows much of what many of us already know from sad experience. Divorced from their patients, concentrating on their business, toying with medical technology, doctors often alienate the sick and frightened with impoverishing bills, high-handed manners and bureaucratic mazes. He lays the blame for these failures on elite training at medical schools, authoritative and paternalistic attitudes, and the advantages of power, prestige and money. Arrogant doctors, he claims, frequently act decisively when hesitation and doubt are best; often they perform no better than quacks and faith-healers.

From greed, doctors tend to over-treat and over-test:

When a surgeon is faced with a choice between doing nothing for which he collects a thirty-dollar consulting fee, and recommending surgery for which he would get three thousand dollars, is he not going to be influenced in the direction of surgery? Only a physician would dare suggest impartiality under the circumstances.

Preston cites studies showing that Americans who come to their doctors for help are twice as likely to undergo surgery as their British cousins. Since the overall mortality rate in the two countries is roughly the same, either Americans are more prone to illness — which seems unlikely — or a vast number of pointless operations are being performed. And because of over-medication and useless therapies, many patients become ill — some even die — *because* they visited their doctor. Had some stayed away, allowing healing to take its natural course, they would have remained alive and well, unharmed by iatrogenic disease.

Preston's most cutting attack is against doctors who practice medicine as if it were a guessing game, relying solely on their subjective clinical judgement rather than proven scientific methods. The random, controlled trial is the only sound basis on which doctors can claim effective therapies; yet physicians continue to walk down unscientific paths. Preston shows that medicine without science is doomed to repeat its recent and dangerous errors — radical mastectomy, chelation treatment for arteriosclerosis, pure oxygen for premature babies, electroshock therapy for mental illness. Ironically, it was under the banner of improved scientific standards that physicians campaigned for and ultimately won control over the practice of medicine in the United States.

Preston draws parallels between American doctors today and ancient healers, taking us back to tribal communities where the shaman and the primitive

medicine-man dispensed their arts. He tries to show us how in fundamental ways medical practice has remained unchanged for millennia, that the basic needs of patients and the old ways of practising medicine emerge as a natural consequence of social interaction. He argues that unless patients rise up and fight for their rights, doctors, who always have had the upper hand, will continue to exercise unreasonable control.

His argument from history is, however, sketchy and flawed. Anthropologists tell us that healers in tribal societies did quite different things in different cultures for their sick and dying. Levi-Strauss reports that the "shaman" in one society may act the self-conscious charlatan, reaping the rewards of power and prestige; in another, the healer may be someone who merely happens to know tribal lore. Moving on, Preston also forgets to inform us that mediaeval barbers and surgeons practised their trades in the same guild. Nor does he tell us that, until modern times, surgeons and chefs were in the same fraternity in the British navy. What bound the cooks, barbers and surgeons was not their natural social standing, but their knives and scissors — or perhaps the fact that food, hair and disease have always been and will always be with us.

By making us believe that the doctor-patient relationship is a natural feature of the human landscape, Preston stumbles and falls into the same myth-making apparatus that holds us all in its thrall. While he makes some recommendations for reform, he fails to see how he has adopted the same myths that have helped American doctors secure their power and privilege.

Preston would have been a better guide had he not turned the pages of his history book so far back. In the nineteenth century, American physicians practised alongside homeopaths, midwives and assorted fakers and purveyors of patent medicine. It was not until the early days of this century that doctors, under the political leadership of the American Medical Association, codified their position and under the slogans of standardization and faith in science, systematically excluded all but those who emerged from approved medical schools and who gained access to practice by passing state licensing exams (also controlled by the AMA and its local societies). By the 1920s, the AMA had established a complete monopoly over the practice of medicine in the United States.

In *The Clay Pedestal*, Thomas Preston has tackled a subject which demands close and critical scrutiny. But an analysis of medicine that merely reflects what patients already feel about their doctors, and neglects to tell them how doctors gained their power, says a lot about what we already know and nothing about what we do not. □

Robert Ubell is the American Publisher of Nature.

Around the world of mental illness

G. Morris Carstairs

Psychiatry Around The Globe: A Transcultural View. By Julian Leff. Pp.204. ISBN 0-8247-1532-2. (Dekker: 1981.) SwFr. 58.

THIS conspectus of mental illness "around the globe" has many of the attributes of those luxury cruises in which travellers combine the pleasures of exotic sight-seeing with those of being instructed by experts about what they have seen. Dr Leff is an admirable tour guide. He handles statistics with respect but is always ready to point out their limitations as well as their contributions to knowledge. As befits an epidemiologist, he first scrutinizes the methodology of any survey and only then discusses its findings.

His global *tour de l'horizon* is clearly presented, under four major headings: (i) Do psychiatric conditions look the same in different countries? (ii) Do psychiatric conditions have the same frequency in different countries? (iii) Are psychiatric conditions treated differently in different countries? (iv) Do psychiatric conditions have a different course in different countries?

These four parts are followed by a final section discussing the mental health of immigrants in general, and of West Indian and Asian immigrants in Britain in particular. He gives us interesting summaries of research findings in each of these areas of inquiry. From start to finish one basic question keeps recurring, namely: if multiple culture-specific factors enter into the perception, treatment and outcome of psychiatric illnesses, then how meaningful is it to make cross-cultural comparisons of these illnesses?

There is now ample evidence that both organic and functional psychoses can be identified in every society; but even in respect of these major illnesses we have to proceed cautiously in interpreting symptoms in the same way as we would do in the West. As Leff puts it: "It is clear that delusions and hallucinations, the main symptoms of psychotic illness, can only be judged as present in relation to the patient's cultural milieu". This still leaves us uncertain as to just how such symptoms are to be evaluated.

A large part of this book is thus devoted to demonstrating ways in which the procedures of psychiatric diagnosis are being made more valid and reliable than they have been in the past. An interesting example is the US/UK Project which had its origin in an observation of the biostatistician, Morton Kramer. He pointed out that, if their respective data were to be believed, then the first admission rates for manic depressive psychoses in the age group 55 to 64 were twenty times higher in Britain than in the USA. Could this really be true?

In order to put the question to the test it was necessary to construct an objective, systematic interview schedule and to train a team of six psychiatrists (four British and two American) in its use. The instrument chosen was Professor John Wing's Present State Examination (PSE). The team used it to ascertain the "project diagnosis" for consecutive series of patients admitted to hospitals serving London and New York. When they did so, the previously wide discrepancies in diagnoses became very much reduced. Analysis of the results showed that hospital doctors in New York were accustomed to use a much wider concept of schizophrenia than that used in London, as a result of which they included under the label of schizophrenia many patients who would be diagnosed as manic-depressive in London.

This was the first international use of the PSE. Its next major deployment was in the WHO-sponsored International Pilot Study of Schizophrenia (IPSS) in which teams in nine countries (the USA, Britain, Colombia, Czechoslovakia, Denmark, India, Nigeria, Russia and Taiwan) collaborated in collecting cohorts of patients who met certain defined criteria, assessing them in terms of the PSE and then following them up for a period of five years. Here the PSE could be used in its original form only in Britain and the USA; in each of the seven other countries it had to be translated into the local language.

This international study (which still continues) has produced some interesting findings; among them, the demonstration that patients with specified "core symptoms" of schizophrenia could be found in every country; and also that the final diagnostic judgements (whose criteria were not laid down) were fairly uniform in seven countries but widely divergent in the other two, namely the USA and Russia. In the former, a cultural emphasis on Freudian concepts led to a wide concept of schizophrenia so that many cases were included under this heading which would not be so regarded in the other countries. In Russia, diagnostic practice was — and still is — dominated by Professor Snezhnevsky's theory that schizophrenia presents in three forms, much attention being paid to the postulated course of the disease. Many patients were diagnosed as schizophrenic because although they did not yet show major symptoms it was predicted that they would eventually do so.

Dr Leff is candid about the large obstacles which hinder cross-cultural agreement about certain of the symptoms of mental illness, particularly in the neuroses. He devotes a chapter to discussing "The language of emotion", pointing out that there are some languages in which there are no single terms for "depression" or "anxiety". In translating

an instrument such as the PSE into Yoruba, for example, recourse has to be made to metaphors in common use such as "the heart is weak", for depression, and "the heart is not at rest", for anxiety. One cannot be sure that these phrases correspond at all precisely with the English words.

Not only are some languages richer than others in the language of emotion: even within cultures members of underprivileged groups may have less well-developed abilities to recognize or to express nuances of feeling (as Basil Bernstein has demonstrated in British society).

Dr Leff is dedicated to the use of standardized instruments, accompanied by glossaries and sets of instructions as the best way to enable different observers to recognize and record symptoms in the same way. Usually he remains keenly aware of the difficulties which still remain, but very occasionally he lapses as when, after completing a review of a number of surveys of psychiatric illness in Asian countries he draws attention to the wide differences in the prevalence rates for neurosis, and comments: "We can infer that the very highest rates found in the Asian surveys are likely to represent the most realistic estimates of the true prevalence of the neuroses".

Normally, he would be the first to question the concept of the "true prevalence" of neuroses. He knows that minor neurotic symptoms are very common in normal subjects. Determining at what levels of frequency and severity these symptoms should constitute an illness is a judgemental decision. Until some less subjective indicator can be found, the prevalence of neuroses can only be measured in terms of stated, hopefully relevant, operational criteria. As yet there is no such thing as the "true prevalence" of neuroses — and perhaps never will be.

The chapters on the mental health of immigrants to the USA and to Britain are interesting not only for their findings (that, in general, immigrants show a high rate of incidence of schizophrenia but that Asian immigrants attend their general practitioners for minor complaints less often than do their British neighbours), but also because they show that useful cross-cultural research can now be carried out within the British Isles and USA. One factor, however, appears to have been overlooked. This is Srole *et al.*'s report, *Mental Health in the Metropolis: The Midtown Manhattan Study* (McGraw-Hill, 1962), which found very high rates of psychiatric disorders among immigrants — but only among poor immigrants. The much smaller number of rich immigrants did not experience an excess of these disorders.

The final chapter tells us, in grim detail, about the many handicaps which an Asian immigrant physician has to contend with in trying to advance his career in a Western country: these are economic, linguistic,

social, educational as well as professional. Having taken a special interest in the welfare of overseas postgraduate students at the Institute of Psychiatry, London, Dr Leff knows what he is talking about; but here he seems to present a totally gloomy picture. Every experience of the immigrant doctor is a negative one, culminating for some in a neurotic or psychotic breakdown, with all their disastrous implications. Is there no gleam of light at all in the experience of our immigrant trainees? Surely we can all remember some exceptions, men and women who have earned the friendship and respect of their British or North American colleagues and have embarked upon an academic or clinical career. Admittedly, they are in the minority. This chapter, although tangential to the rest of the book, will serve a useful purpose by reminding us about the difficult lot of our overseas trainees. □

G. Morris Carstairs is currently a Fellow of the Woodrow Wilson Center, Washington DC, and recently spent three years in India as a visiting professor of psychiatry.

Art of insects

J.L. Cloudsley-Thompson

Insects Etc.: An Anthology of Arthropods Featuring A Bounty of Beetles. Pp.108. ISBN 0-933920-25-3. (Hudson Hills: 1981.) \$50.

WHAT is there in common between Edgar Allan Poe, Jonathan Swift and Raoul Hausmann? The answer, immediately revealed by perusal of the literary anthology with which P.A. Gette enlivens this unusual book, is that they were not good observers of insects. Like Salvador Dali's pre-1950 ants, Hausmann's houseflies had only four legs, Swift's were endowed with stings and Poe confused bugs with beetles. Far more observant were Lewis Carroll and Vladimir Nabokov and, of course, Réaumur, Thoreau and Jean-Henri Fabre. Buffon was evidently a careless writer: "The insect is a small animal

without blood. The latter [sic] can be divided into large and small kinds".

Literary activities, like art, are not necessarily marred by inaccuracy such as this. Less excusable, however, are the errors of nomenclature and spelling apparent in the names of the various arthropods portrayed. For instance, *Argiope bruennichi* (mis-spelled) is an impressive, striped, orb-web spinner, but it is not the zebra spider. This name has always been reserved for *Salticus scenicus* the little jumping spider, common on walls and fences throughout Europe, Asia and America. Again, the scorpion is not *Buthus occitanus* as stated but, unmistakably, *Scorpio maurus*, the classical "scorpio" of antiquity.

The 34 dazzling, large-format colour plates are taken from photo-realistic paintings by Bernard Durin, without doubt an unusually gifted illustrator in the great tradition of Audubon, Edward Lear and Louis Agassiz Fluertes. Of 33 species illustrated (there is a close-up of the head of *Zonocerus variegatus* in addition to a painting of the whole grasshopper), 16 are beetles, 5 grasshoppers, 6 Hemiptera, 3 Hymenoptera, a single butterfly (the peacock), a spider and a scorpion. The species depicted are, as might be expected, large, brightly coloured and sometimes bizarre. The paintings illustrating them are breathtaking. Durin is not only a meticulous observer but he really succeeds in creating the illusion of iridescence in his beetles and in making insects' wings look transparent. A commentary on each of the arthropods illustrated, with information about various aspects of its behaviour, life cycle, food, enemies and folklore, is provided by Gerhard Scherer.

Zoologists are often so familiar with the subjects of their research that they overlook obvious questions posed by them. What are the functions of the extraordinary colours of some insects, why do others have such long legs or antennae, what are the advantages and drawbacks of their shapes and sizes? It might be worthwhile taking a second, thoughtful look at the subjects of Durin's paintings. □

J.L. Cloudsley-Thompson is Professor of Zoology at Birkbeck College, University of London.



Pycna antinorii, the African cicada, painted by Bernard Durin. Reproduction in black and white and reduction in size do scant justice to the picture.

REVIEW ARTICLE

Mantle metasomatism—continuing chemical change within the Earth

D. K. Bailey

Department of Geology, University of Reading, Whiteknights, Reading RG6 2AB, UK

Mantle fragments carried up from as deep as 200 km by high-velocity eruptions show metasomatic introduction of volatile bearing minerals, before melting. The broad spatial and temporal distribution of alkaline magmatism shows that transfer of lithophile elements to the outer Earth by rising volatiles, is a long-term and continuing process in planetary evolution.

METASOMATISM is a chemical change, whereby a pre-existing mineral or rock is converted to another composition. It usually refers to a solid-state transformation, with material transfer through a vapour or fluid, without melting. Paradoxically it has always found easiest acceptance when applied to low-temperature and -pressure phenomena, for example, the detailed replacement of calcium carbonate fossils by iron sulphide, or the substitution of plant tissue by silica in fossil wood. Mining geologists have long accepted metasomatism as an important process, giving rise to a whole class of metasomatic or 'replacement' ore bodies, spectacular examples of which are wholesale replacement of limestone beds by metal sulphides¹. Obviously, some evidence of the original material must remain as proof of metasomatism, and with higher temperatures and pressures this becomes more vulnerable. Controversy is most likely when deep metasomatism is proposed as the origin for a relatively common rock such as granite²: the end product is chemically unexceptional, and the issue complicated by rock recrystallization (metamorphism) at high temperatures and pressures. Even so it should be noted that most metamorphic changes are not isochemical because the mineral reactions involve at least volatile losses or gains³.

On a limited scale, metasomatic replacement in crustal rocks is well documented⁴⁻⁶, new minerals formed including feldspar, tourmaline, kaolin, micas, sulphides, serpentine, carbonates, calc silicates, alkali amphiboles and pyroxenes. All cases involve volatile or mobile elements from the range K, Na, C, H, F, Cl, S and B, and the metasomatic process operates by fluid infiltration or diffusion. The efficacy of processes of this kind is demonstrated by the oxygen isotope equilibration between large rock masses and circulating groundwaters⁷. Nonetheless, as a major rock-forming process in the crust, metasomatism may still be controversial, so that mantle metasomatism (where the evidence is fragmentary) might seem esoteric. The concept arose from the need to integrate geological, petrological and geophysical observations on cratonic alkaline magmatism^{8,9}. This concept was supported by the fact that crustal rocks intruded by alkaline magmas provide perhaps the best examples of alkali metasomatism, now known as 'fentization' from the type area at Fen (Southern Norway)¹⁰. Subsequently, mantle fragments brought up by alkaline volcanism were found to be metasomatized, such that if melted they would yield the compositions observed as lavas¹¹. Support came from chemical studies on alkaline lavas from oceanic islands, indicating variations in the source mantle that would be explicable by alkali enrichment before melting¹².

Composition of the mantle

The mantle constitutes the major part of the Earth (84% of the volume and 68% of the mass) and is the chemical and energy reservoir driving crustal and lithogenic processes. As most of the information about the mantle, and all below 250 km,

is derived from geophysical observations, the picture that emerges is inevitably smoothed by the limited resolving power of the indirect observations. Furthermore, remote physical measurements are unlikely to discern anything less than gross variations in chemistry (and even then the domains of chemical variations would have to be very large). Consequently there is some danger of treating the Earth's mantle as if it were chemically homogeneous or (perhaps even more dangerous) rendered homogeneous soon after accretion of the proto-Earth. For convenience it is customary to regard the bulk composition of the mantle as being essentially peridotitic (dominantly magnesium-rich silicates, olivines and pyroxenes) similar to the silicate fractions of stony meteorites¹³, particularly the abundant chondrites. Meteorite compositions are widely used as reference standards for minor (trace) elements and isotopes, but the once popular chondrite Earth model has (with some exceptions¹⁴) largely fallen into disfavour¹⁵. An internationally agreed standard for bulk mantle chemistry would have value in general geochemical comparisons, but any chemical reference standards are liable to abuse in the sense that they can all too readily assume the status of a real piece of mantle from which specific surface rocks were derived. Garnet peridotite, as brought to the Earth's surface through diamond pipes, is the preferred choice for a generalized bulk composition^{13,16} and although Ringwood prefers, as a geochemical device, to use a hypothetical material, pyrolite, (a chemical mix of peridotite and basalt) he notes that garnet peridotite is "rather similar in composition"¹⁵. Beneath this apparent unanimity there is diversity, ranging from the "bulk composition of the mantle is very uncertain"¹⁶ to there is "major element homogeneity throughout the mantle source regions" of basaltic magmas¹⁵. It is, however, largely from the difficulty in explaining the variations in chemistry between magmas that concepts of mantle heterogeneity have grown¹⁷. Some types of heterogeneity go beyond those expected from irreversible changes caused by previous expulsion of melts to the surface, and seem to require additions or mixing of sources earlier in the Earth's history.

Arguments for metasomatic changes in mantle composition fall into two categories—those deduced from chemical characteristics of lavas, and those where there is observed replacement of an earlier mineralogy. Both need to be seen in a geological context. Direct observation has claim to precedence but it is convenient to deal with lava chemistry first, because the case has been developed in the less complex framework of oceanic volcanism, and because the two lines of investigation become most compelling in unison.

Chemical variations in the oceanic mantle

All the evidence for mantle metasomatism is provided by igneous activity, especially mafic (basic) and ultramafic magmatism for which there is abundant evidence of origin at least in the upper mantle, and maybe deeper. As liquid samples from

the mantle, the magmas themselves may be expected to carry some imprint of the source chemistry. Deciphering the imprint depends on simplifying assumptions, some of which may be easily acceptable, but all are debatable. Basic lavas erupted within the ocean basins are generally believed to be uncontaminated by continental crust and surface materials. Because of their voluminous and widespread eruption, mid-ocean ridge basalts (MORB) are taken to be the products of high degrees of melting of the source mantle, and to have many chemical characteristics that are unmodified by subsequent crystal fractionation. Isotope ratios of heavy elements such as Sr, for example, are held to be immune¹⁸⁻²⁰, but all these views have been challenged²¹. Most geochemists working on MORB, however, are satisfied with the assumptions as a working hypothesis, which allow the use of element concentrations, ratios and isotopes to characterize the chemistry of the source mantle. Trace elements that because of their ionic size and charge cannot be accommodated in the crystal lattices of the main peridotite (mantle) minerals are said to be 'incompatible' and hence strongly partitioned into any melt that may form. Using crystal-melt partition coefficients it is possible to calculate expected fractionation patterns of various elements, for different amounts of melting of plausible combinations of mantle minerals. The calculated patterns may then be compared with those observed in MORB. The first attempt gave the best fit for the MORB incompatible element patterns at 15-30% melting of a model mantle peridotite, whereas alkali basalts from oceanic islands seemed to be restricted to as little as 3-7% melting if they were to give the required patterns²². Also the higher concentrations of a range of large ion elements in alkali basalts, compared with MORB, required that the source mantle was richer in these elements than the MORB source²². Gast postulated that the mantle supplying alkali basalts was more primitive, and hence the MORB source had lost some of its large ion elements in an earlier melting episode. He described the MORB source as "depleted" with respect to the supposedly primitive alkali basalt source. Behind this notion of depletion was the knowledge that there is commonly insufficient of the large ion radioactive isotope ⁸⁷Rb in MORB to support the observed amounts of the daughter isotope ⁸⁷Sr, which is most readily explained by some Rb loss in the earlier history of the mantle source. Subsequently, it became clear that by the same token ocean island alkali basalts had too much Rb for their content of ⁸⁷Sr, and this was consistent with their source having been 'enriched' in large ions some time before eruption²³. In current usage, 'enriched' often means just 'more abundant',

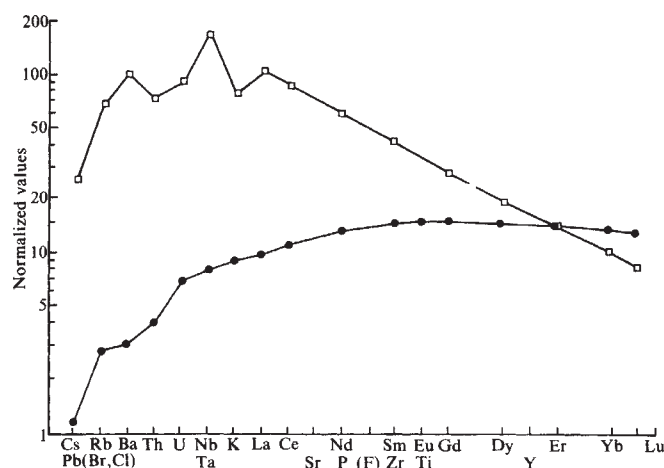


Fig. 1 Comparison of normalized abundance patterns of incompatible elements from alkali basalt (□) (oceanic islands) and mid-ocean ridge tholeiite basalt (●). Adapted from ref. 25. Refractory lithophile elements such as rare earths, Ti, Zr, U and Th normalized against chondritic meteorites. For other normalizing values see ref. 25. Highly alkaline lavas such as nephelinite will in general plot at higher levels than alkali basalt.

and may even start to assume the connotation of 'metasomatized'. Before describing something as 'enriched' or 'depleted' its previous condition must be known. This leads back to the most difficult assumptions, concerning the origin and evolution of the planet. Typical are discussions about rare earth elements in lavas, in which the abundances are compared by normalizing against a chondrite meteorite standard. For example, the lower concentrations of the light rare earths (La, Ce, Pr, Nd) in MORB, causing a downward dip in their distribution pattern when normalized against chondrite abundances, may be said to be due to an earlier depletion in the mantle source. Taken alone this conclusion can only be based on an expectation that an undepleted mantle source would provide liquids with light rare earth distributions similar to, or higher than in chondrites: either way, assumptions have been made about the source, and a real (but unseen) segment of mantle has been invested with some of the attributes of an arbitrary geochemical standard. This can lead to bewildering mixtures of hypotheses and observations, but without making any assumptions about the previous history of the mantle, it is still possible to deduce from the chemistry of lavas that their sources were different at the time of melting. More and more geochemists have favoured heterogeneity in the mantle to account for the differences between fresh lavas of similar major element chemistry, when these have been erupted contemporaneously in the same region¹⁷. In favourable circumstances it may be possible to conclude in favour of heterogeneity even on the basis of major element ratios in lavas, for example, Mg/Fe (ref. 24), but most cases are based on the range of incompatible elements depicted in Fig. 1.

If chemical heterogeneity in the oceanic mantle is accepted, the debate can turn about the nature and extent of the variations, and their cause and timing. Often the most that can be deduced from lava chemistry is that the sources were richer or poorer in particular elements by comparison among the samples, or against some chosen standard. By adding isotope systematics, it can be argued that the observed lava isotope patterns cannot be explained by undisturbed isotopic evolution in the source throughout the entire history of the Earth (Fig. 2). This proposition is made more plausible by the difficulty in conceiving of a segment of mantle that could have remained totally isolated for 4,600 Myr, but note that the isotope systematics are based on meteorite abundances and must rest on assumptions about isotopic distributions in the primordial Earth. If it is accepted that the systematics define the expected behaviour in a closed system within the Earth, the departures from the geochron indicate that the isotopic development of a particular piece of mantle has been disturbed. Depletion then usually means Rb deficiency, and is almost always ascribed to earlier episodes of melt extraction from primordial, primitive, or fertile mantle. Similar deductions have been made for Nd/Sm (ref. 27), Th/Pb, U/Pb and Pb isotope distributions: diagrams similar in principle to Fig. 2 may also be constructed, but Nd/Sm data are less abundant, and the complexities of the Pb systems require more extensive coverage for satisfactory presentation²⁵. Taken in concert the isotope patterns in a lava can make an impressive case for a complex history of its mantle source²⁵.

Enrichment (such as Rb excess) implies source mixing, the simplest interpretation being that two sources have been mixed, but the nature of the mixing, the number of mixes, and the real number of sources, are matters of conjecture. It might be that melts from deeper sources have become lodged in and contaminated higher mantle levels, which are then subsequently remelted^{12,28-31} or a variety of melts have mingled in their passage to the surface, maybe as a complex continuum. Alternatively, fluids from another mantle source have contaminated melts at higher levels³² or metasomatized the mantle through which they pass³³. The lava erupted at the surface may be the product of any combination of such processes, or others as yet unknown. But all that the 'disturbed' isotope patterns in lavas reveal is that the mantle sources have not had a simple history. Careful integration of different chemical evidence, such as

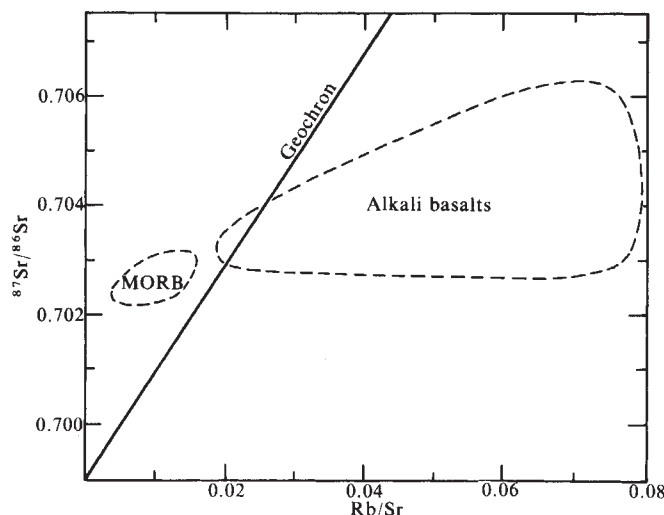


Fig. 2 Rb-Sr isochron diagram showing fields of alkali basalts and MORB. The 4,600 Myr isochron (geochron) indicates the distribution that would be expected from isolated (closed) mantle domains that had the same primordial $^{87}\text{Sr}/^{86}\text{Sr}$ as basaltic achondrite meteorites (BABI), but different Rb/Sr when the Earth was formed. Compositions left of the line have Rb insufficient for, and those to the right Rb in excess of, the amount necessary to support their observed ^{87}Sr , if the geochron prescription is correct. The simplest explanation is that the Rb-Sr systems in the mantle sources have not remained undisturbed, suggesting loss of Rb from the MORB sources and gain by the island volcano sources. From refs 12, 26.

combined isotopes, trace elements and element ratios, can be developed into a forceful argument for source enrichment before melting³⁰. With improved geochemical understanding, and more information, it may be possible to discriminate in favour of enrichment by volatile activity, and metasomatism may then be the most plausible process; but the ultimate proof of metasomatism can be imprinted only in the mantle solids.

Volcanological evidence

Mafic and ultramafic lavas form the bulk of the material for geochemical conclusions in favour of mantle metasomatism, and it is therefore ironic that often the geology or volcanology of the specimens is accorded little or no weight in the argument. Those igneous rocks with the highest contents of incompatible elements have long been referred to by petrologists as 'alkaline', so that abundance of K and Rb, two of the prime incompatible elements in modern discussions, is already implicit. Alkaline magmatism, oceanic and continental, is characteristically expressed in central volcanoes and igneous complexes, giving rise to fragmental, explosive and sometimes high velocity eruptions³⁴. It is these last that carry up ultramafic fragments (nodules), providing solid samples of the deepest available mantle. Furthermore, fast eruption allows little opportunity for modification of the melt as it rises to the surface, and provides some of the best samples of mantle melting. Other incompatible elements are thus manifest in alkaline igneous activity, namely the gases which are the cause and driving force of high speed eruption. Furthermore, the gases in this magmatism are distinctive, especially in their contents of CO_2 , Cl and F (refs 34–36). Therefore richness in alkalis (and other incompatible elements) and gases is no coincidence⁸, and was a prime reason for proposing mantle metasomatism as an essential part of alkaline igneous activity⁹. This does not say that the volcanological evidence is anything other than circumstantial (as indeed is the geochemistry of lava samples), simply that in considering the case for mantle metasomatism all the evidence should be weighed.

Against this volcanological background, the petrographical evidence of metasomatism in mantle nodules brought to the surface by igneous eruptions should be viewed.

Petrographical evidence

To know that metasomatism has happened, the evidence must be seen in the rock, and solid samples of mantle are available only as volcanic xenoliths, and peridotite masses in fold mountain belts. The latter may have been emplaced as relatively cool tectonic slices, or hot intrusions, forced through the Earth's crust. Some contain veins of less refractory minerals, and veined peridotite has been held up as one exemplar of 'enriched' mantle that might yield alkali basalts by preferential melting of the vein material^{30,37,38}. The veins have been interpreted as formed by melt injection, or segregation. Evidence of metasomatism while the peridotite was still in the mantle is not yet available, and may not be easy to distinguish from low-pressure changes after emplacement in the crust. Hence the only petrographical evidence of metasomatism is in mantle nodules of volcanic derivation. Even here careful observation is necessary to eliminate the possibility that the changes did not result from reactions between a peridotite fragment and the high temperature, volatile-rich melt which brought it to the surface.

Peridotite nodules consist dominantly of olivine and orthopyroxene with, or without, small amounts of clinopyroxene, and/or spinel, and garnet. It is these last three minerals especially that contain the Ca, Al and Na, and are consequently important contributors to basaltic liquids, the most common surface products of mantle melting. Spinel peridotites are typically carried up in alkali basalt (basanite) eruptions, while garnet-bearing varieties are much more restricted, and largely confined to strongly silica-unsaturated volcanics, such as melilitites, and especially kimberlites. All these nodule carriers are notably rich in incompatible elements, but even tholeiitic basalts have K and H_2O contents that are hard to account for by realistic degrees of melting of peridotite unless it contains small amounts of an additional phase, such as amphibole³⁹ or phlogopite mica⁴⁰. Experimental studies^{40–43} indicated that these minerals had appropriate stabilities for existence in the upper mantle, but early reports of amphibole and mica in peridotite nodules were circumspect, either because the minerals seemed secondary (and formed outside mantle conditions)⁴⁴ or because the peridotite nodules might be magmatic crystal accumulations (and not accidental inclusions picked up by the magma)⁴⁵. But evidence was mounting in favour of most peridotite nodules being mantle samples^{46,47}, and textural studies soon left little doubt that in some peridotites, phlogopite⁴⁸ and amphibole⁴⁹ had grown in equilibrium. These hydrous minerals were seen as either an integral part of the mineralogy of the mantle sampled by the eruption^{48,49}, or formed as a result of igneous veining around a magma body in the mantle⁵⁰. In one case it was suggested that direct melting of amphibole peridotite could have produced the nepheline melt that brought the fragments to the surface⁴⁹.

Mica and amphibole are the two most characteristic metasomatic minerals in peridotite nodules, the process being recognized by intricate partial replacement of the earlier minerals. New minerals invade the old in delicate lobes and embayments, until the stage may be reached where only scattered (but undisturbed) residuals remain: the rock is pervaded by infiltration ('seepage') of new material along grain boundaries, and even along cleavage planes in the existing minerals¹¹. Clinopyroxene also forms metasomatically, its composition being distinctly different from any that may be present in the original peridotite^{51,52}. A suite of samples can show all stages from incipient metasomatism of peridotite through to complete transformation into mica-amphibole-clinopyroxenite. This latter rock has been found containing residuals of peridotite still intact¹¹. The minerals most susceptible to replacement are garnet, spinel and orthopyroxene, followed by olivine. Other new minerals introduced are carbonates, phosphate (apatite), and titanium and iron oxides and sulphides^{11,53,54}. By its very nature, metasomatism must yield varied end products, but the process may add K, Ti, Al, Fe, Mn, Ca, Na, H, C, S, P, Rb, Sr, Y, Zr, Nb, Ba and La (ref. 11), in different amounts. This is an impressive

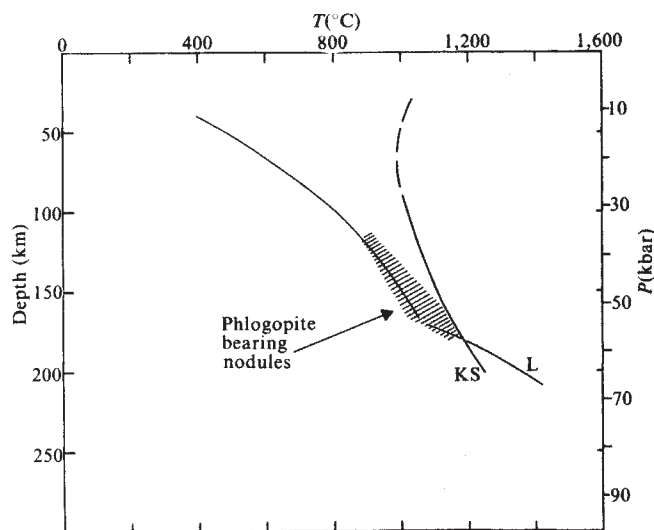


Fig. 3 Distribution of phlogopite-bearing mantle nodules from Lesotho kimberlites in relation to the geothermal gradient (L), and kimberlite beginning-of-melting (KS). Phlogopite has formed well outside the conditions where any melt can exist, and cannot be attributed to the action of magma. On the contrary, the melting curve may be the upper boundary for phlogopite survival in retrieved mantle samples. The distribution also suggests a maximum depth of ~200 km for metasomatized samples. Adapted from ref. 58.

array of incompatible elements, and the concentrations may also be remarkable: one opaque phase has been reported⁵⁴ with 9.1% BaO and 7.4% ZrO₂. In addition to the pervasive alteration within the body of the rock, metasomatized peridotites are commonly cut by veins^{11,53,54} containing assemblages of metasomatic minerals. These veins are truncated against lava at nodule margins, indicating that the nodules are fragments of larger masses disintegrated by eruption. In some cases there are small amounts of glass in the veins, with every sign that the liquid was in equilibrium with the vein minerals. This liquid, indicating the melting threshold, is in stark contrast with the host lava which corrodes and devolatilizes the hydrous minerals to clinopyroxene + olivine ($\pm$ spinel) + glass^{11,55}. Furthermore, there are chemical gradients perpendicular to the veins within the nodules but not against the nodule margins⁵⁶. Thus, a range of observations rule out metasomatism of the nodules by the carrier melt at low pressure.

Not uncommonly the mica and amphibole in peridotite nodules is deformed^{54,57} in keeping with a time lag between its formation and nodule eruption. Furthermore, comparing mica bearing nodules with the mantle solidus on a *PT* diagram⁵⁸ (Fig. 3) shows that they equilibrated in conditions where no liquid could exist. Isotopic analyses of Sr in the metasomatic minerals and host nodules prove them to be out of equilibrium with the enclosing kimberlite, and indicate a significantly earlier metasomatic age⁵⁹. In other examples, although the metasomatized nodules and the recently erupted host lavas are in an isotopic harmony, the isotopic evidence requires the melts to be fusion products of mantle that had been metasomatized in an earlier event⁶⁰. Other studies reveal complex histories of changing chemical heterogeneity in the continental mantle, some of which were metasomatic, dating back to 3,000 Myr (refs 61, 62).

Conditions of metasomatism

Considerations of mechanisms (infiltration and diffusion) and the characteristics of the metasomatizing agents, are still in their infancy. Studies of the gaseous components of metasomatic amphiboles and micas indicate that the aqueous fluids coexisting with those amphiboles and micas brought up by alkali basalts were diluted by CO₂, O, F and Cl, while in samples from kimberlites all the H₂O may have been fractionated into

hydrous minerals⁵⁶. Experimental studies have already shown that fluids in equilibrium with synthetic phlogopite and olivine must contain considerable dissolved K, Al, and Si (ref. 3), and the agent of mantle metasomatism (whatever its precise nature) was clearly bringing in these and other large ion lithophile elements from greater depth. The ultimate, deep mantle source of the fluids is inaccessible, but the emission of gases from volcanoes is evidence of continual Earth degassing, and the key role of CO₂ in alkaline magmatism is of long standing³. Repetition of alkaline igneous activity in continental cratons predicates a persistent, or renewable, source of volatiles in the mantle^{34,36}. The same volatiles must be essential ingredients of the metasomatic process.

Figure 4 shows some experimentally determined mineral stabilities, and indicates that the most abundant metasomatic minerals in ultramafic nodules could have been formed by fluids moving through the mantle along typical geothermal gradients. Figure 3 shows this to be borne out by the metasomatized nodules in kimberlite pipes, and also indicates the expected upper boundary of the process—the onset of melting where the local geothermal gradients intersect the vapour-saturated solidus in the mantle. On low geothermal gradients, and with slow release of fluids from the deep mantle, metasomatism will continue and presumably affect a wider region, over a long period of time. But given a sufficient supply of fluid the channels will become saturated with metasomatic minerals, and their temperatures will be raised, so that ultimately melting and the generation of alkaline magmas may be expected⁸. Such a regime explains the broad range of nodule types showing all degrees of metasomatism, and with some showing signs of melt formation. Upward leakage of volatiles can thus provide the means of chemical enrichment, and the trigger for igneous activity^{8,58}. A mechanism for volatile release and channeling through the lithosphere is shown in Fig. 5.

Volume relations

In fine detail the metasomatism in ultramafic nodules is non-disruptive, with undisturbed residual patches of the original host. At this scale part of the replacement may be equal-volume,

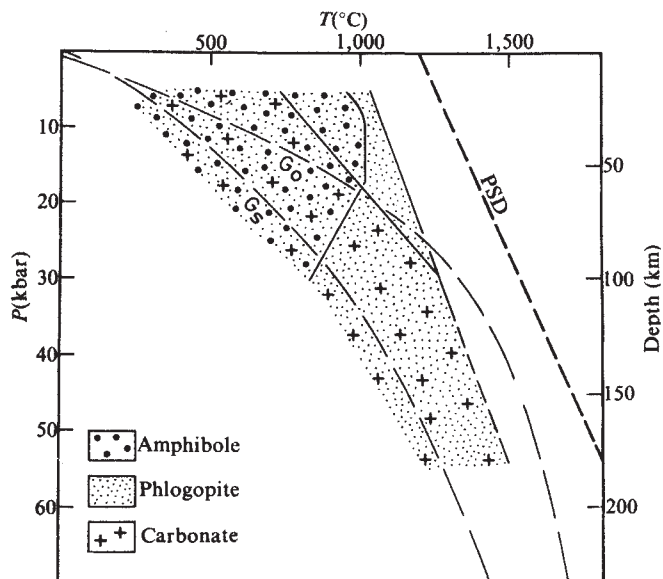


Fig. 4 Experimentally determined stabilities for possible metasomatic minerals, the boundaries (solid lines) indicating upper limits. Overlap of the stability fields cannot be used to infer necessary coexistence of minerals, which depends on factors such as bulk mantle chemistry, presence or absence of vapour (fluid), and fluid composition. Extensive mantle stability regions are indicated for continents (geotherm G_c) where alkali rich and carbonate magmatism are strongly developed. Mineral stability ranges are much more restricted along typical oceanic gradients (G_o). PSD (heavy broken line) is the vapour absent peridotite solidus. Adapted from ref. 63. Experimental sources, refs 40, 43, 64, 65.

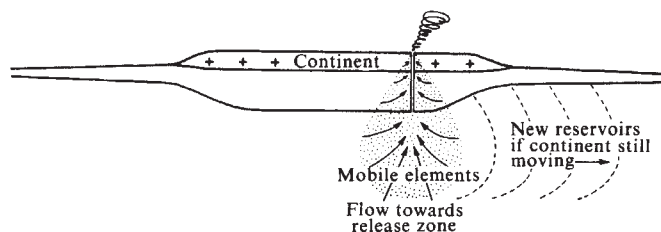


Fig. 5 Opening of a zone of weakness in the lithosphere allows volatile release to the surface. The escape channel effectively acts as a volatile and heat focus, leading to metasomatism and melting along the pathways. Adapted from ref. 34.

but it is almost impossible to rule out the possibility of rock distension along grain boundaries. Complete equal-volume metasomatism would require totally free exchange between the rock and the transporting medium, so that part of the original composition can be removed, as suggested in nodules from the Bultfontein kimberlite⁵⁴. In mantle conditions the bulk of any fluid entering a region where metasomatic minerals were stable would be consumed in carbonation and hydration reactions. Expansion is then unavoidable, making fluid escape zones self-extending by what amounts to metasomatic wedging. Consistent with this is the common observation of veining in metasomatized nodules. Once a crack system is open in the host rock, chemical exchange between fluid and solid must become easier, evidenced by the new minerals being relatively homogeneous. Complex zoning and other heterogeneities are common, however, in metasomatic minerals, indicating a spectrum of replacement conditions. Metasomatic veining itself requires some opening up of the host rock, even though most of the minerals along the veins may have formed by replacement. A parallel exists in alkaline metasomatism (finitization) around alkaline complexes in the crust, where the outstanding physical characteristic is extensive shattering and veining of the host rock, often with little sign of deformation, or even block rotation. Mantle metasomatism requires that a pre-existing volume of dense peridotite has been changed to one that is larger and less dense, which explains the remarkable correlation of alkaline magmatism with uplift of the cratons (coupled with geophysical evidence of a lower density mantle)^{9,11}. Uplift by this process has the advantage of long survival, in contrast to uplifts of purely thermal origin. In fact the metasomatic uplift can be removed only by surface erosion or by subsequent thermal decomposition in the mantle: the latter may explain part of the subsidence observed in continental rifts when the axial igneous activity becomes established³⁶.

Is metasomatism widespread or local?

First descriptions of metasomatized peridotites indicated two different modes for the process. On one hand the metasomatized nodules formed only a fraction of the total mantle sample in a kimberlite vent, and were attributed to local metasomatism by magma before eruption⁵³. At the other extreme, throughout two whole volcanic provinces large proportions of the ultramafic nodules were metasomites, sometimes to the exclusion of unaltered peridotites¹¹. Substantial parts of the sampled mantle in south-west Uganda and west Eifel (Germany) consist of alkali clinopyroxenite, and in the Uganda samples garnet and orthopyroxene are absent, and olivine not abundant. The lavas in these (and like) provinces have always been burdened by a plethora of names, giving little or no indication of their mineralogical affinities and obscuring the fact that in essence they are fine-grained clinopyroxenites. Their chemistry is appropriate to their being partial melts from alkali clinopyroxenite mantle: K-rich lavas in Uganda where the nodules are rich in mica, and sodi-potassic lavas in the Eifel where amphibole is more abundant in the nodules^{11,66}. A parallel case can be made for kimberlite, which in its earliest description was called mica peridotite. Nodule suites in which large proportions of the samples contain amphibole and mica are not uncommon, and there seems to

be a spectrum of conditions from extensively metasomatized mantle to a patchy or network distribution. Where metasomatism is extensive (and intensive) the lavas look like daughter products, and there is growing support for the view that metasomatism is the precursor of alkaline magmatism^{54,60,67}.

Even a volcanic province such as south-west Uganda, however, is only a tiny mantle sample, and the need now is to see metasomatism in the context of the mantle as a whole. Clearly material is being introduced into the upper mantle from below, but is it only a local effect⁶⁸⁻⁷¹ and therefore of relatively small significance in mantle processes, or is it more widespread^{9,11,61}? Work on nodules from different kimberlites and of different ages across the southern African craton has demonstrated the use of coordinating different lines of enquiry⁶¹. Mineral textures show metasomatic replacement in nodules from many kimberlites, and a detailed Sr isotopic study of nodules from one kimberlite suggests they have a distinctly greater age than that of the host kimberlite (90 Myr), the metasomatic event being tentatively related to the Karroo magmatism (190 Myr) that affected wide areas of Africa and the rest of Gondwanaland^{59,72}. Metasomatism in nodules from the Precambrian Premier kimberlite indicates an earlier manifestation of the process in the same region⁷³. New Nd isotope studies are indicating two different episodes of metasomatism in nodules from the same kimberlite (Bultfontein) (A. J. Erlank, personal communication). More and more kimberlite pipes are being discovered in southern Africa⁶⁸, and although the volume of kimberlite in each case is tiny, the southern African craton has clearly been peppered with kimberlite eruptions, providing widespread sampling of the underlying mantle. Proportions of metasomatized nodules vary^{53,54} so the metasomatism is patchy in mantle sampled by kimberlite; but the time relations, and the spatial distribution, require that the process, although variable in intensity, is general below the southern African craton. Similar patterns emerge elsewhere, and the world-wide occurrence of kimberlite and alkaline magmatism bringing up mantle xenoliths containing amphibole, mica, and apatite suggests that metasomatism of the mantle, especially the subcontinental mantle, is widespread. The multiplicity of alkaline magmatism across the continents throughout their history indicates that metasomatism is not just an anomalous feature in an otherwise barren mantle. The age ranges, from Precambrian through to present, indicate a continual process, and the persistent association of metasomatized nodules with alkaline magmatism must surely mean that the metasomatizing fluids and the magmatism repeatedly exploit, as channels, long-lived weaknesses in the lithospheric plates⁷⁴.

Wider considerations

Geochemistry, and observational and experimental petrology show that the mantle is heterogeneous. Much of the Earth's thermal activity is expressed in submarine eruption at mid-ocean ridges, and this has lent credence to the view that the continental upper mantle is largely depleted in its low melting constituents, and is barren, infertile, or effectively dead. Active volcanoes in continental interiors indicate a counter view, but their relative paucity perhaps causes them to be set aside as anomalies⁷⁵, or the progeny of mythical plumes. It is easy to overlook the fact that the cratons are the oldest parts of the Earth's surface, and since they became stable they have been repeatedly perforated, across their breadth, by alkaline igneous activity. Even tholeiitic flood lavas erupted on a regional scale may show significantly higher levels of incompatible elements than could be expected from a highly refractory residual mantle⁶¹, and some have chemical features close to that of the postulated bulk Earth⁷⁵. The magmatism is a constant reminder that development of the continental mantle is far from complete. Repeated and widespread metasomatism of the mantle lithosphere would account for this activity, and would also be expected to contribute to continental growth by additions to the base of the crust⁹. Continental permanence and constancy of crustal thickness could be maintained in the face of the fact

that erosion now exposes Archaean deep crust at the surface. If the continental mantle is not exhausted, it raises the possibility that the special characteristics of the ancient cratons are inherited from original accretionary inhomogeneities in the formation of the Earth³⁴. Undoubtedly, the continental lithosphere still has challenges that can engage the scientific collaboration envisaged in the International Lithosphere Programme.

Perhaps a start could be made by dismissing the view that the old continental heartlands had completed their evolution before ours began.

I thank A. J. Erlank, G. N. Hanson, F. E. Lloyd, M. A. Menzies and S.-S. Sun for support in the preparation stages, and for their stimulating work in this field; also A. J. Erlank and F. E. Lloyd read the first draft and gave helpful comments.

1. Barnes, H. L. *Geochemistry of Hydrothermal Ore Deposits* 2nd edn (Wiley, New York, 1979).
2. Read, H. H. *The Granite Controversy* (Murby, London, 1957).
3. Bailey, D. K. in *The Evolution of the Crystalline Rocks* (eds Bailey, D. K. & Macdonald, R.) 416–469 (Academic, London, 1976).
4. Shand, S. J. in *Eruptive Rocks* (Murby, London, 1972).
5. Turner, F. J. & Verhoogen, J. *Igneous and Metamorphic Petrology* (McGraw-Hill, New York, 1960).
6. Nockolds, S. R., Knox, R. W. O'B. & Chinner, G. A. (Cambridge University Press, 1978).
7. Taylor, H. P. *J. geol. Soc. Lond.* **133**, 509–558 (1977).
8. Bailey, D. K. *Geol. J. Spec. Iss. No. 2*, 177–186 (1970).
9. Bailey, D. K. *J. Earth Sci., Leeds* **8**, 225–239 (1972).
10. McKie, D. in *Carbonates* (eds Tuttle, O. F. & Gittins, J.) 261–294 (Wiley, New York, 1966).
11. Lloyd, F. E. & Bailey, D. K. *Phys. Chem. Earth* **9**, 389–416 (1975).
12. Sun, S. S. & Hanson, G. N. *Contr. Miner. Petrol.* **52**, 77–106 (1975).
13. Yoder, H. S. Jr *Generation of Basaltic Magma* (National Academy of Sciences, Washington DC, 1976).
14. Bougault, H., Joron, J. L. & Treuil, M. *Phil. Trans. R. Soc. A297*, 203–213 (1980).
15. Ringwood, A. E. *Composition and Petrology of the Earth's Mantle* (McGraw-Hill, New York, 1975).
16. Smith, J. V. *Proc. 8th Lunar Sci. Conf.* 333–369 (1977).
17. Bailey, D. K., Tarney, J. & Dunham, K. C. *Phil. Trans. R. Soc. A297*, 135–493 (1980).
18. Faure, G. & Powell, J. L. *Strontium Isotope Geology* (Springer, Berlin, 1972).
19. Faure, G. *Principles of Isotope Geology* (Wiley, New York, 1977).
20. Cox, K. G., Bell, J. D. & Pankhurst, R. J. *The Interpretation of Igneous Rocks* (Allen and Unwin, London, 1979).
21. O'Hara, M. J. *Phil. Trans. R. Soc. A297*, 215–227 (1980).
22. Gast, P. W. *Geochim. cosmochim. Acta* **32**, 1057–1086 (1968).
23. Sun, S. S. & Hanson, G. N. *Geology* **297**, 297–302 (1975).
24. Langmuir, C. H. & Hanson, G. N. *Phil. Trans. R. Soc. A297*, 383–407 (1980).
25. Sun, S.-S. *Phil. Trans. R. Soc. A297*, 409–445 (1980).
26. Brooks, C., Hart, S. R., Hofmann, A. & James, D. E. *Earth planet. Sci. Lett.* **32**, 51–61 (1976).
27. Menzies, M. & Murthy, V. R. *Earth planet. Sci. Lett.* **46**, 323–334 (1980).
28. Varne, R. & Graham, A. L. *Earth planet. Sci. Lett.* **13**, 11–18 (1971).
29. Green, D. H. *Phil. Trans. R. Soc. A268*, 707–725 (1971).
30. Tarney, J., Wood, D. A., Saunders, A. D., Cann, J. R. & Varet, J. *Phil. Trans. R. Soc. A297*, 179–202 (1980).
31. Frey, F. A. & Green, D. H. *Geochim. cosmochim. Acta* **38**, 1023–1059 (1974).
32. Schilling, J.-G., Bergeron, M. B. & Evans, R. *Phil. Trans. R. Soc. A297*, 147–178 (1980).
33. Hanson, G. N. *A. Rev. Earth planet. Sci.* **8**, 371–406 (1980).
34. Bailey, D. K. *Phil. Trans. R. Soc. A297*, 309–322 (1980).
35. Bailey, D. K. in *The Alkaline Rocks* (ed. Sorensen, H.) 53–66 (Wiley, New York, 1974).
36. Bailey, D. K. in *Petrology and Geochemistry of Continental Rifts* (eds Neumann, E.-R. & Ramberg, I. B.) 1–13 (Reidel, Dordrecht, 1978).
37. Sun, S. S. & Hanson, G. N. *Geology* **4**, 626–631 (1976).
38. Hanson, G. N. *J. geol. Soc.* **134**, 235–253 (1977).
39. Oxburgh, E. R. *Geol. Mag.* **101**, 1–19 (1964).
40. Yoder, H. E. Jr & Kushiro, I. *Am. J. Sci.* **267A**, 558–582 (1969).
41. Boyd, F. R. in *Researches in Geochemistry* Vol. 1 (ed. Abelson, P. H.) 377–396 (Wiley, New York, 1959).
42. Kushiro, I., Syono, Y. & Akimoto, S. *Earth planet. Sci. Lett.* **3**, 97–203 (1968).
43. Kushiro, I. *Yb. Carnegie Instn Wash.* **68**, 245–247 (1970).
44. Dawson, J. B. & Powell, D. G. *Contr. Miner. Petrol.* **22**, 233–7 (1969).
45. O'Hara, M. J. in *Ultramafic and Related Rocks* (ed. Wyllie, P. J.) (Wiley, New York, 1967).
46. Leggo, P. J. & Hutchison, R. *Earth planet. Sci. Lett.* **5**, 71 (1968).
47. Kleeman, J. D., Green, D. H. & Lovering, J. F. *Earth planet. Sci. Lett.* **5**, 449 (1969).
48. Dawson, J. B., Powell, D. G. & Reid, A. M. *J. Petrol.* **11**, 519–48 (1970).
49. Varne, R. *J. Petrol.* **9**, 169–190 (1968).
50. Wilshire, H. G. & Trask, N. J. *Am. Miner.* **56**, 240–251 (1971).
51. Wass, S. Y., Henderson, P. & Elliott, C. J. *Phil. Trans. R. Soc. A297*, 333–346 (1980).
52. Lloyd, F. E. *Miner. Mag.* **44**, 315–323 (1981).
53. Harte, B., Cox, K. G. & Gurney, J. J. *Phys. Chem. Earth* **9**, 477–506 (1975).
54. Erlank, A. J. & Rickard, R. S. *Extended Abstr. 2nd int. Kimberlite Conf.*, Santa Fe (1977).
55. Francis, D. M. *J. Petrol.* **17**, 357–378 (1976).
56. Boettcher, A. L. & O'Neil, J. R. *Am. J. Sci.* **280-A**, 594–621 (1980).
57. Boyd, F. T. & Nixon, P. H. *Geochim. cosmochim. Acta* **42**, 1367–1382 (1978).
58. Bailey, D. K. *Miner. Mag.* **43**, 695–699 (1980).
59. Erlank, A. J. & Shimizu, N. *Extended Abstr., 2nd int. Kimberlite Conf.*, Santa Fe (1977).
60. Menzies, M. & Murthy, V. R. *Am. J. Sci.* **280A**, 622–638 (1980).
61. Erlank, A. J., Allsopp, H. L., Duncan, A. R. & Bristow, J. W. *Phil. Trans. R. Soc. A297*, 295–307 (1980).
62. Menzies, M. & Murthy, V. R. *Nature* **283**, 634–636 (1980).
63. Bailey, D. K. *9th Tomkeiff Memorial Lecture*, Newcastle upon Tyne, 1–41 (University Geology Department, 1979).
64. Eggler, D. H. & Holloway, J. R. *Proc. American geophys. Un.: Chapman Conf. on Partial Melting in the Earth's Upper Mantle*, 15–36 (1977).
65. Wyllie, P. J. *J. geol. Soc. Lond.* **134**, 215–234 (1977).
66. Duda, A. & Schmincke, H.-U. *Neues Jb. Miner. Abh.* **132**, 1–33 (1978).
67. Wass, S. Y. & Rogers, N. W. *Geochim. cosmochim. Acta* **44**, 1811–1823 (1980).
68. Dawson, J. B. *Kimberlites and Their Xenoliths* (Springer, Berlin, 1980).
69. Wilshire, H. G., Pike, J. E. N., Meyer, C. E. & Schwarzmann, E. C. *Am. J. Sci.* **280A**, 576–593 (1980).
70. Wyllie, P. J. *J. geophys. Res.* **85**, 6902–6910 (1980).
71. Gurney, J. J. & Harte, B. *Phil. Trans. R. Soc. A297*, 273–293 (1980).
72. Kramers, J. D. *Earth planet. Sci. Lett.* **42**, 58–70 (1979).
73. Erlank, A. J. *et al. IUGG XVII General Assembly*, Canberra (1979).
74. Bailey, D. K. *J. geol. Soc. Lond.* **133**, 103–106 (1977).
75. O'Nions, R. K., Evensen, N. M. & Hamilton, P. J. *Phil. Trans. R. Soc. A297*, 479–493 (1980).

ARTICLES

Resolving crystallographically distinct tetrahedral sites in silicalite and ZSM-5 by solid-state NMR

C. A. Fyfe^{*†}, G. C. Gobbi[†], J. Klinowski[‡], J. M. Thomas^{*†} & S. Ramdas[‡]

[†] Guelph-Waterloo Centre for Graduate Work in Chemistry, Department of Chemistry, University of Guelph, Ontario, Canada N1G 2W1

[‡] Department of Physical Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EP, UK

'Magic-angle-spinning' NMR has revealed fresh insights into the atomic structure of the polymorph of silica, silicalite, which has exceptional absorptive properties. There is evidence that highly crystalline variants of silicalite contain 24 distinct tetrahedral locations, populated principally by Si⁴⁺ ions, and that, contrary to what was originally thought, aluminium is also present, in at least two kinds of environments, in the tetrahedral framework. The silicalite structure is shown to be essentially indistinguishable from that of the siliceous zeolite catalyst ZSM-5.

THE new, thermally stable, crystalline phase of silica called silicalite described by Flanigen *et al.*¹, possesses remarkable sorptive properties that stem partly from its exceptionally large internal volume (33% porosity) and partly from its three-

dimensional system of intersecting channels¹. It contains five-membered rings in basic repeats (Fig. 1a) which are so arranged as to form a system of channels made up of 10-membered ring openings (Fig. 1b). Silicalite is hydrophobic, and it removes from water, by preferential adsorption, alkanols, certain alkanes and some other organic entities. It has been argued¹ that

* To whom correspondence should be addressed.

"...unlike aluminium-containing zeolites, silicalite has no cation-exchange properties... silicalite has no aluminium and no cations in its structure". The original patent² asserts that such aluminium as may be adventitiously present is in the form of Al_2O_3 . Although the details of its crystal structure remain to be fully clarified, it is established^{1,3-5} that, topologically, silicalite is closely similar if not identical to the shape-selective catalyst ZSM-5, which because it selectively converts methanol to petrol, is a key component in the production of fuel from coal ($\text{C} + \text{H}_2\text{O}(\text{steam}) \rightarrow \text{CO} + \text{H}_2$; $\text{CO} + 2\text{H}_2 \rightarrow \text{CH}_3\text{OH}$). There have been claims⁴ that silicalite is an end member of the ZSM-5 substitutional series, which embraces Si/Al ratios from 10 to 4,000. But because it is extremely difficult, if not impossible⁵, to ascertain by X-ray crystallography whether aluminium in very low concentration is present substitutionally in the corner-sharing tetrahedral (SiO_4) array that constitutes ZSM-5 and silicalite, there is still considerable debate as to the distinction between these two materials. We now show, by magic-angle spinning⁶ NMR, that: (1) several crystallographically distinct sites for Si atoms in the lattice may be directly detected by ^{29}Si NMR; (2) the aluminium in silicalite is readily detectable by ^{27}Al NMR; (3) the Si/Al ratio may be straightforwardly estimated; (4) all the aluminium is tetrahedrally coordinated to oxygens; (5) there are at least two distinct types of tetrahedral site occupied by the aluminium. We further show that there are 24 distinct crystallographic tetrahedral sites occupied by either silicon or aluminium in silicalite/ZSM-5, and that its

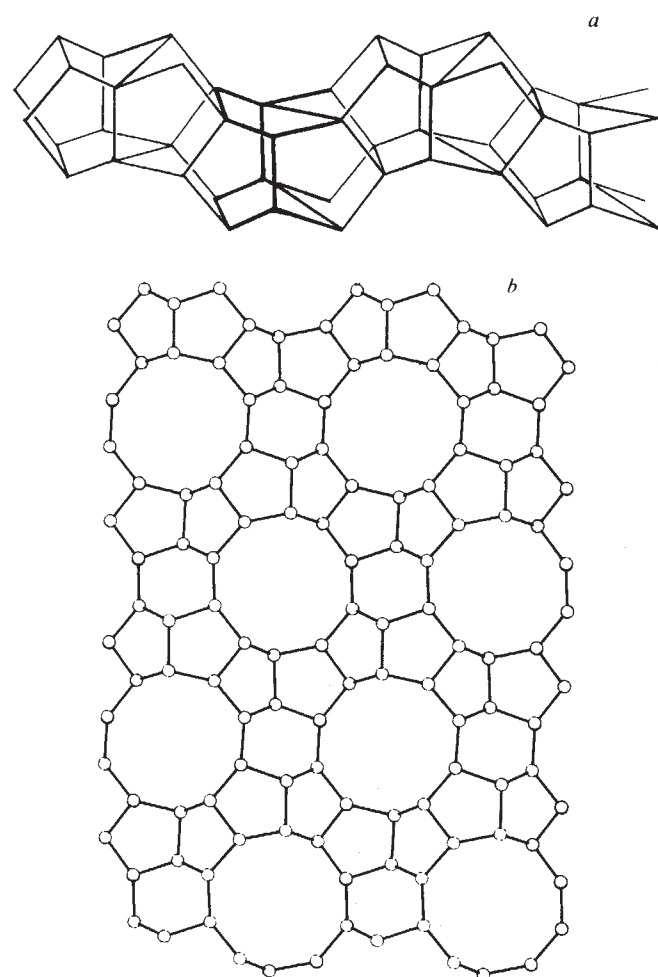


Fig. 1 The structure of silicalite. *a*, Secondary building units (indicated by bold lines) each composed of 12 Si atoms are linked into chains, one of which is shown in the *c* direction. *b*, The chains are interlinked to form a three-dimensional framework in which there are 10-membered ring openings (~5.5 Å diameter) running along the [010] direction. In this portion of the *ac* structural projection, O denotes a tetrahedral site.

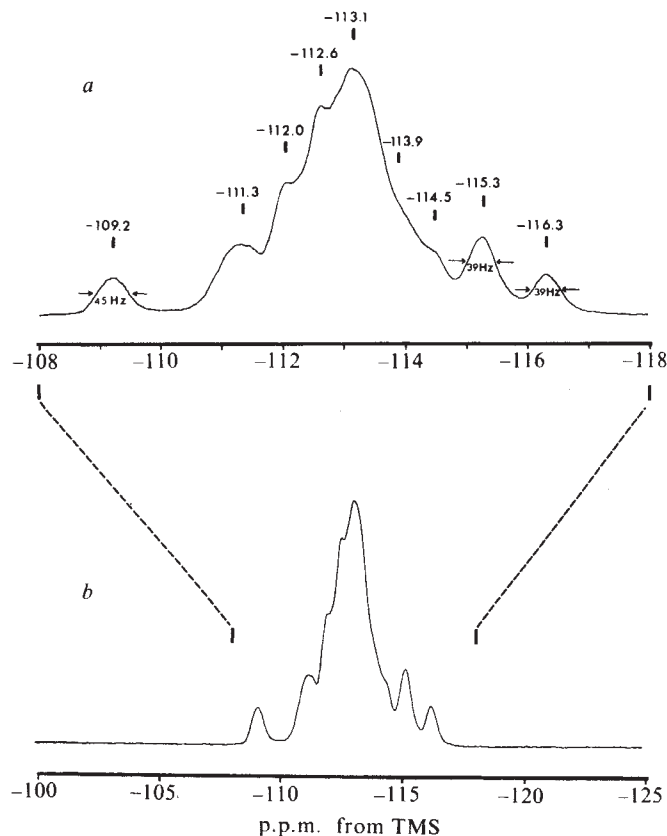


Fig. 2 *a, b*, High-resolution ^{29}Si MASNMR spectrum of silicalite at 79.80 MHz. 6,550 free induction decays were accumulated; repetition time 5 s. Chemical shifts are given in p.p.m. from TMS.

space group has lower symmetry than the currently accepted Pnma ³. We conclude that, structurally, silicalite and ZSM-5 are essentially indistinguishable.

Structure determination

High-resolution solid-state ^{29}Si NMR with magic-angle spinning (MASNMR) is capable of distinguishing between the five possible $\text{Si}(n\text{Al})$ structural building blocks in zeolitic aluminosilicates, where $n = 0$ to 4 is the number of tetrahedral Al atoms joined, by oxygen bridges, to the central Si atom⁷⁻¹⁰. In the spectra of crystalline synthetic faujasites (zeolites X and Y) for example there are up to five well-resolved signals falling within distinct chemical shift ranges; thus the value of n can be assigned to each signal, aiding structural elucidation¹¹⁻¹³.

By using ^{29}Si MASNMR at very high magnetic field¹⁴ on a sample of silicalite of very high purity and crystallinity prepared exactly as given in example 4 of ref. 2 but afterwards calcined in air at 500 °C for 16 h, we have obtained spectra of such high resolution that we can resolve separate ^{29}Si signals attributable to several crystallographically non-equivalent tetrahedral Si atoms, all joined, by oxygen bridges (denoted by straight lines in Fig. 1*a, b*), to four other Si atoms. These spectra, together with the corresponding ^{27}Al MASNMR spectra, provide new insights into the structure of silicalite.

Results

Figure 2 shows the ^{29}Si spectrum recorded at 79.80 MHz without using any resolution-enhancement techniques. The sharpness of the spectrum is exceptional, with FWHM of the resolved peaks of ~40 Hz (~0.5 p.p.m.) which is superior to that so far observed in any other zeolite, even at this very high magnetic field. The complete spread of all the components of the spectrum (chemical shifts are given from external Me_4Si) is ~6 p.p.m. The chemical shift range of all the peaks (-112 ± 3 p.p.m.) is characteristic of $\text{Si}(0\text{Al})$ (that is $\text{Si}(4\text{Si})$) groupings

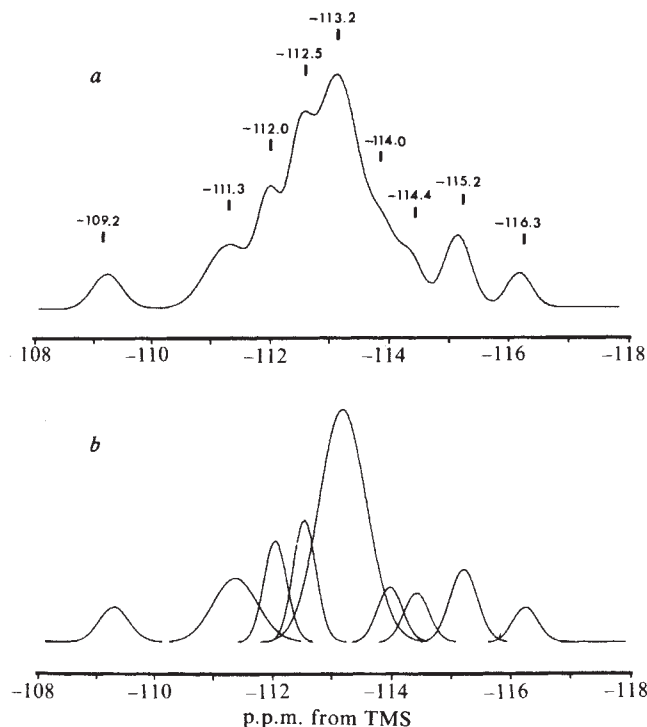


Fig. 3 The ^{29}Si MASNMR spectrum given in Fig. 2 can be computer-simulated using nine gaussian-shaped peaks, shown individually in *b*; their sum, simulating the spectrum, is given in *a*. The areas of the peaks in *b* are, from left to right, in the ratio 0.98:2.70:2.19:2.63:10.35:1.30:1.16:1.87:0.82.

in highly siliceous materials⁷⁻¹⁰. The observed multiplicity arises from the crystallographically non-equivalent tetrahedral environments of Si(4 Si) silicons. The spectrum may be simulated by a minimum of nine gaussian signals, the intensities (peak areas) of which are approximately in the ratio 1:3:2:3:10:1:1:2:1 (see Fig. 3). Using the intensities of the well-resolved lowest and highest field signals as a base unit of one (that is we proceed on the plausible assumption that these well-resolved peaks each refer to one distinct tetrahedral site), the total intensity of the peaks in the Si(4 Si) multiplet is found to be approximately 24. This result suggests a space group which requires 24-non equivalent Si sites in the repeat unit of the structure. It signifies that $\text{Pn}2_1/a$ or $\text{P}2_1/n$ (see below) are more appropriate than Pnma^3 which has only 12 distinct tetrahedral sites. The relative intensities of the signals do not suffer perceptible change over a 10-fold increase in the delay time of the experiment, thereby indicating that they are quantitatively reliable. At this stage, it is premature to attempt to assign individual ^{29}Si peaks to specific sites, particularly because

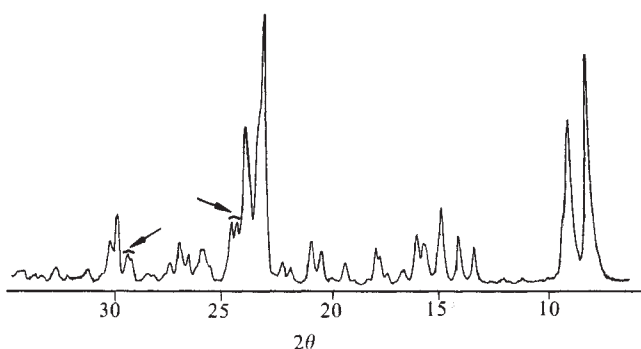


Fig. 4 Powder diffraction pattern of silicalite obtained on a Philips vertical goniometer using Cu $K\alpha$ radiation selected by a curved graphite monochromator in the diffracted beam. Peak doublets are arrowed.

interatomic distances in silicalite are not sufficiently well established (see footnote 11 of ref. 5).

Figure 4 shows a powder X-ray diffraction pattern of silicalite (the same sample as used in the MASNMR measurements shown in Fig. 2). Note the presence of peak doublets (marked with arrows) for 2θ values of 24.4° and 29.2° . Pseudosymmetry is common in zeolite structures and monoclinic as well as orthorhombic symmetry has been observed in ZSM-5-type materials^{15,16}. Wu *et al.*¹⁵ simulated peak splitting by the introduction of monoclinic lattice parameters $a = 20.17 \text{ \AA}$, $b = 19.93 \text{ \AA}$, $c = 13.42 \text{ \AA}$ and $\alpha = 90.64^\circ$, instead of the orthorhombic values of $20.07 \text{ \AA}$, $19.92 \text{ \AA}$, $13.42 \text{ \AA}$ and 90° , respectively. This small change in the unit cell causes the symmetry

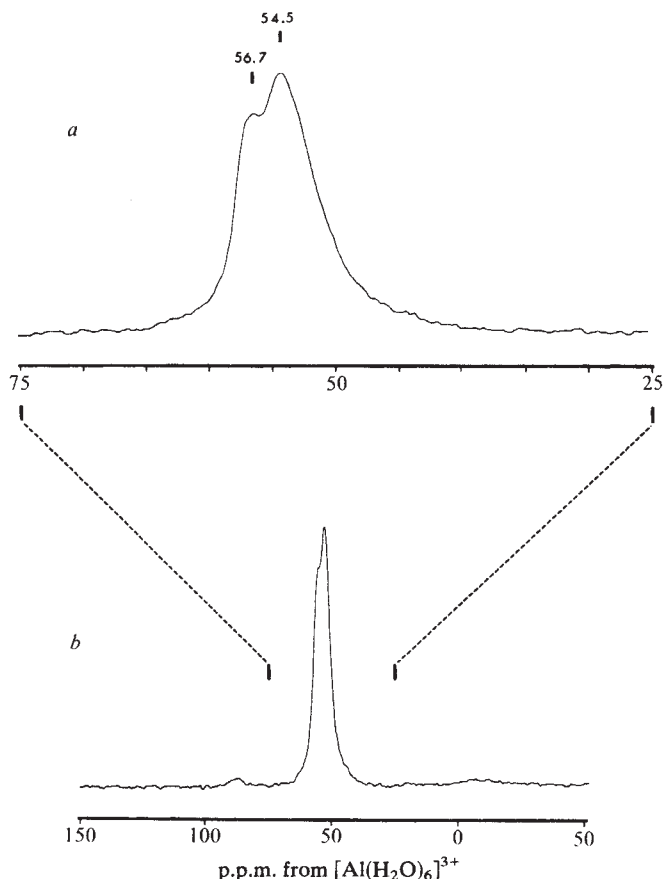


Fig. 5 High-resolution ^{27}Al MASNMR spectrum of silicalite at 104.22 MHz. 176,214 free induction decays were accumulated; repetition time 0.1 s. Chemical shifts are given in p.p.m. from $[\text{Al}(\text{H}_2\text{O})_6]^{3+}$. Note the absence of octahedrally coordinated aluminium in the sample.

elements such as the mirror planes parallel to (010) bisecting the zigzag-shaped channels to disappear; there are thus 24 (rather than 12) crystallographically non-equivalent Si atoms. The probable monoclinic space group for ZSM-5-type materials is, therefore, thought¹⁵ to be $\text{P}2_1/n$. Change of symmetry occurs quite readily for ZSM-5 materials: indeed, a commercially available sample of silicalite exhibited an X-ray pattern identical to that shown in Fig. 4.

Discussion

The ^{27}Al MASNMR spectra show a rather broad absorption centred at ~ 55.6 p.p.m. with respect to external $[\text{Al}(\text{H}_2\text{O})_6]^{3+}$, clearly characteristic of tetrahedral coordination¹⁷, both in this and in all other silicalite samples, which we have examined. In the sample the ^{29}Si spectrum of which is given in Fig. 2, the peak shows fine structure (Fig. 5) due to at least two components (at 54.5 and 56.7 p.p.m.) indicating the presence of crystallographically non-equivalent sites for tetrahedral aluminium. This agrees with the ^{29}Si spectrum in Fig. 2 and indicates that

the aluminium is tetrahedrally incorporated into the framework. Because of the 100% isotopic abundance of ^{27}Al and its very short spin-lattice relaxation time, even traces of aluminium are detectable by this sensitive technique. The aluminium in silicalite comes from impurities in the materials used in synthesis, and their level varies from sample to sample. The amount of Al present is small, but in all the samples of silicalite we have examined Al is tetrahedral and not present as occluded Al_2O_3 impurity² as the various forms of Al_2O_3 which we have investigated for comparison show an intense NMR peak at 9 ± 1 p.p.m. (reflecting octahedral coordination of Al) and in some also a small peak at 72 ± 3 p.p.m. due to tetrahedrally coordinated¹⁷ Al, in agreement with structural data¹⁸ on γ -alumina. A quantitative determination by ^{27}Al MASNMR of Al concentration in the sample of silicalite was carried out by adding a known amount of a reference compound containing octahedrally coordinated Al and comparing the intensities of the two signals. This gives a Si/Al ratio $\geq 1,000$. The aluminium content of the commercial silicalite sample was found to be higher (Si/Al ≥ 200) and a small broad peak at -104 p.p.m. corresponding to Si(1Al) (that is Si(0Al) (0Si)₃ group) is clearly observed in the ^{29}Si MASNMR spectrum.

We have found general correspondence between the ^{29}Si MASNMR spectra of silicalite on the one hand, and those for

a variety of ZSM-5 zeolite preparations on the other. The latter show in the ^{29}Si MASNMR spectra an intense peak at about -112 p.p.m. with some indications of fine structure and a much smaller peak at approximately -105 p.p.m. corresponding to Si(1Al); and in the ^{27}Al spectrum a single peak at ~ 55 p.p.m., identical to that found at essentially this position in silicalite. The ^{29}Si spectra of ZSM-5 zeolites can, to a first approximation, be simulated by applying a large line-broadening function to the silicalite spectrum of Fig. 2, thus confirming the close structural relationship between the two materials.

This work demonstrates the power of MASNMR as a structural probe for a large class of shape-selective aluminosilicate sorbents and catalysts.

We acknowledge support from the Universities of Cambridge and Guelph and from BP Research Centre, Sunbury. C.A.F. acknowledges the financial support of the Natural Sciences and Engineering Research Council (NSERC) of Canada in the form of an operating grant, G.C.G. the award of a NSERC Graduate Scholarship. We thank Dr D. Stewart, Dr B. Nay and Dr M. Barlow for making available the special preparation of highly crystalline silicalite. The NMR spectra were obtained at the South-West Ontario High Field NMR Facility (Dr R. E. Lenkinski, manager) funded by a Major Installation grant from NSERC.

Received 4 February; accepted 19 February 1982.

1. Flanigen, E. M. *et al.* *Nature* **271**, 512 (1978).
2. Grose, R. W. & Flanigen, E. M. U.S. Patent 4,061,724 (1977).
3. Kokotailo, G. T., Lawton, S. L., Olson, D. H. & Meier, W. M. *Nature* **272**, 437 (1978).
4. Olson, D. H., Haag, W. O. & Lago, R. M. *J. Catal.* **61**, 390 (1980).
5. Olson, D. H., Kokotailo, G. T., Lawton, S. L. & Meier, W. M. *J. phys. chem.* **85**, 2238 (1981).
6. Andrew, E. R. *Int. Rev. Phys. Chem.* **1**, 195 (1981).
7. Lippmaa, E., Mägi, M., Samoson, A., Engelhardt, G. & Grimmer, A.-R. *J. Am. chem. Soc.* **102**, 4889 (1980).
8. Lippmaa, E., Mägi, M., Samoson, A., Tarmak, M. & Engelhardt, G. *J. Am. chem. Soc.* **103**, 4992 (1981).

9. Ramdas, S., Thomas, J. M., Klinowski, J., Fyfe, C. A. & Hartman, J. S. *Nature* **292**, 228 (1981).
10. Klinowski, J., Thomas, J. M., Fyfe, C. A. & Hartman, J. S. *J. phys. Chem.* **85**, 2590 (1981).
11. Engelhardt, G., Lohse, U., Lippmaa, E., Tarmak, M. & Mägi, M. *Z. anorg. allg. Chem.* **482**, 49 (1981).
12. Klinowski, J., Ramdas, S., Thomas, J. M., Fyfe, C. A. & Hartman, J. S. *JCS Faraday Trans II* (in the press).
13. Melchior, M. T., Vaughan, D. E. W. & Jacobson, A. J. *J. Am. chem. Soc.* (in the press).
14. Fyfe, C. A. *et al.* *J. Magn. Res.* (submitted).
15. Wu, E. L., Lawton, S. L., Olson, D. H., Rohrman, A. C. Jr & Kokotailo, G. T. *J. phys. Chem.* **83**, 2777 (1979).
16. Bennett, J. M., Cohen, J. P. & Kirchner, R. M. *Am. Chem. Ass. Meet. Abstr.* **6**, 62 (1979).
17. Fyfe, C. A., Gobbi, G. C., Hartman, J. S., Klinowski, J. & Thomas, J. M. *J. phys. Chem.* (in the press).
18. Wells, A. F. *Structural Inorganic Chemistry* (Oxford University Press, 1962).

Monitoring of structural changes accompanying ultrastabilization of faujasitic zeolite catalysts

J. Klinowski*, J. M. Thomas^{*†}, C. A. Fyfe^{†‡} & G. C. Gobbi[‡]

^{*} Department of Physical Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EP, UK

[†] Guelph-Waterloo Centre for Graduate Work in Chemistry, Department of Chemistry, University of Guelph, Ontario, Canada N1G 2W1

By recording high-resolution solid-state NMR spectra of zeolitic samples spun (at ~ 3 kHz) at the magic angle ($\theta = 54^\circ 44'$ so that $3 \cos^2 \theta - 1 = 0$) to the magnetic field, the chemical identity of the first and second coordination shells surrounding Si may be determined, and a distinction made between octahedrally and tetrahedrally bonded Al. The technique, therefore, enables the atomic course of dealumination, known to be responsible for the stabilization of catalytic zeolites, to be directly determined.

SYNTHETIC analogues of the rare mineral faujasite (Fig. 1) can be prepared in mild laboratory conditions such that materials with a wide range of compositions (zeolites X and Y with Si/Al ratios from 1.18 to 2.75) may be produced^{1,2}. The topology and overall crystallography of the aluminosilicate framework remain constant, but the population of Si^{4+} and Al^{3+} ions in the tetrahedral sites depends on the Si/Al ratio. These zeolitic solids, which are capable of cation exchange, possess large internal areas ($\sim 800 \text{ m}^2 \text{ g}^{-1}$) and volumes ($\sim 50\%$ porosity) and attractive molecular sieving properties. The diameter of the connected intracrystalline cavities varies between ~ 7 and $\sim 13 \text{ \AA}$. It has long been recognized that faujasitic zeolites are good catalysts for the cracking of petroleum; but in their as-prepared condition they proved

inadequate for hydrocracking, mainly because of their excessive acidity. A major turning point in the use of synthetic faujasites in catalytic cracking and hydrocracking came with the demonstration^{3,4} that on heat-treating NH_4^+ -exchanged zeolite Y ultrastable catalysts were produced. These materials can withstand temperatures of $\sim 1,000^\circ \text{C}$, retain their internal surface areas and exhibit desirable hydrocracking activity. Over 200,000 tons of ultrastable zeolites are now produced annually: faujasitic catalysts are the cornerstone of the petroleum industry.

It is known from chemical analysis, and indirectly from the reduction of the unit cell dimension (monitored by X rays) that the Si/Al ratio increases markedly, a fact which both improves the stability and quells the acidity. But what are the atomic rearrangements associated with the process of ultrastabilization? There has been much speculation: it has been surmised

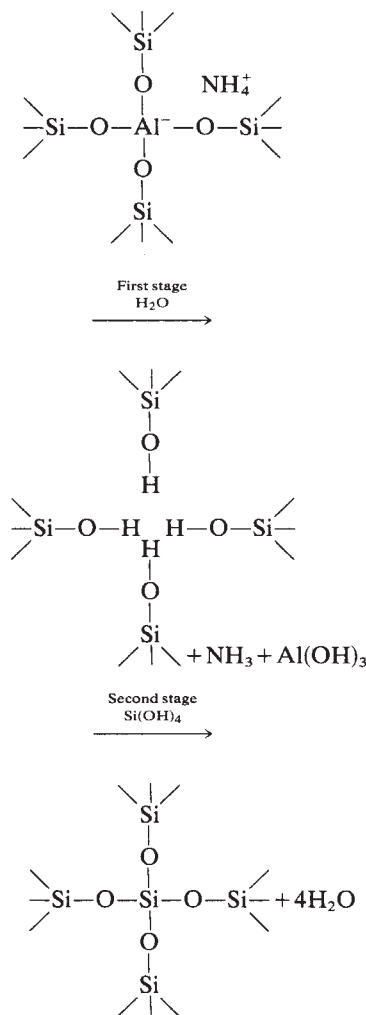
[†] To whom correspondence should be addressed.

that the dealumination gives rise to detrital aluminous entities. In what form is this produced; and which are the groupings in the parent zeolite that are demolished?

Magic-angle spinning NMR

^{29}Si and ^{27}Al magic-angle spinning NMR (MASNMR)⁵⁻¹², is well suited to answer these questions, because it provides direct information pertaining to the immediate local environment of two kinds of atoms. ^{29}Si MASNMR readily and quantitatively identifies the groupings $\text{Si}(\text{O}-\text{Al})_n(\text{O}-\text{Si})_{4-n}$ where $n = 4, 3, 2, 1$ and 0 ; and ^{27}Al MASNMR easily distinguishes^{13,14} between tetrahedrally and octahedrally bound Al^{3+} ions. (For brevity we write $\text{Si}(n\text{Al})$ for the grouping $\text{Si}(\text{O}-\text{Al})_n(\text{O}-\text{Si})_{4-n}$.)

It is proposed¹⁵ that ultrastabilization proceeds as follows:



Thus the framework vacancies created by the removal of aluminium (accompanied by a decrease in ion-exchange capacity) are thought to be subsequently re-occupied by silicon¹⁶ or aluminium¹⁷. The interstitial aluminium can be removed by leaching with acid. Ultrastabilization is facilitated by the presence of water vapour. Materials with very high Si/Al ratios are prepared by repeated ammonium exchange, heat treatment and the removal of interstitial aluminium.

Despite detailed studies of stabilization using various techniques, its mechanism remains unclear, particularly concerning the origin of the silicon required for the proposed second stage. This may come from the surface or amorphous parts of the sample, or from extensive recrystallization. We shall demonstrate that a combination of ^{29}Si and ^{27}Al MASNMR provides new insights into the mechanism of ultrastabilization. In essence, by recording NMR spectra at the magic angle, extremely sharp resonance lines are obtained from solid samples, comparable in width with those normally obtained

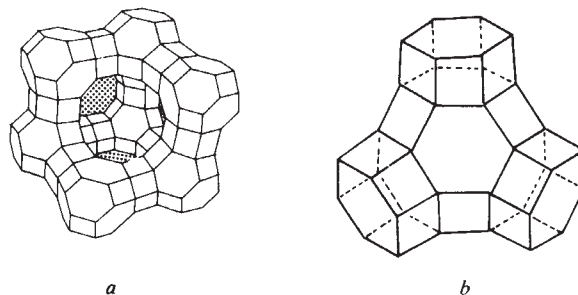


Fig. 1 Some features of the faujasite structure. *a*, The faujasite framework showing 14-hedral cages linked through six-membered rings, and supercages joined through circular windows. Tetrahedral atoms are located at the corners of polygons with oxygen atoms halfway between them. Non-framework cations are not shown. *b*, Tetrahedral linking of 14-hedral cages in faujasite. For clarity, the six-membered ring at the back of the cage is not shown.

from solutions or liquids. Chemical shifts are therefore measurable even in solids.

Experimental

^{29}Si MASNMR spectra were obtained at 79.80 MHz using a Bruker WH-400 narrow-bore spectrometer equipped with a magic-angle spinning attachment¹⁸ of the Andrew-Beams type¹⁹. Spinners, with an internal volume of $\sim 450 \mu\text{l}$, were made of Delrin and spun at a rate of $\sim 3 \text{ kHz}$. Cross-polarization and proton decoupling were not used. Intervals of 5 s were allowed between pulse sequences and 500–5,000 free induction decays were accumulated per sample. All ^{29}Si chemical shifts were measured from external tetramethylsilane, with high field shifts being negative. Experimental spectra have been computer simulated based on gaussian peak shapes, and the relative intensities of $\text{Si}(n\text{Al})$ signals calculated and corrected for the spinning sidebands. ^{27}Al MASNMR spectra were recorded at 104.22 MHz also using a home-made probe. ^{27}Al chemical shifts were recorded with respect to $[\text{Al}(\text{H}_2\text{O})_6]^{3+}$ as an external reference; 3,000 free induction decays were accumulated per sample.

We have examined a series of four samples subjected to different types of treatment. Commercial zeolite Na-Y with $\text{Si}/\text{Al} = 2.61$, unit cell composition $\text{Na}_{53}\text{Al}_{53}\text{Si}_{139}\text{O}_{384} \cdot 240\text{H}_2\text{O}$ (established by X-ray fluorescence) and the (cubic) cell parameter $a = 24.70 \text{ \AA}$ was the starting material. It was 75% exchanged with ammonium sulphate solution yielding $\text{NH}_4\text{-Na-Y}$ zeolite (sample 1). Figure 2*a* shows the ^{29}Si MASNMR spectrum of this material; as expected, it is identical to that for the original Na-Y (not shown). Because in synthetic faujasites the first-order neighbourhood of every Al atom is $\text{Al}(4\text{Si})$, that is the Loewenstein rule strictly applies^{6,11}, the Si/Al ratio of the tetrahedral anionic framework can be calculated^{11,12} from the signal intensities $I_{\text{Si}(n\text{Al})}$ as

$$(\text{Si}/\text{Al})_{\text{NMR}} = \frac{\sum_{n=0}^4 I_{\text{Si}(n\text{Al})}}{\sum_{n=0}^4 0.25nI_{\text{Si}(n\text{Al})}}$$

thus providing a quantitative method of measuring zeolitic composition independently of elemental chemical analysis. Gaussian deconvolution of the ^{29}Si spectrum of sample 1 (Fig. 2*a*) leads to the ratio of intensities of the four peaks of 12.3:36.6:43.8:7.3, corresponding to $(\text{Si}/\text{Al})_{\text{NMR}} = 2.60$, in excellent agreement with the value for the original Na-Y (by XRF and by NMR).

We note that when Si/Al ratios determined for synthetic faujasites by ^{29}Si MASNMR do not coincide with those obtained by chemical methods, it means that non-framework Al or Si (such as detrital silica or adventitiously bound aluminous species) are present. This is clearly not the case in sample 1, the ^{27}Al spectrum of which is given in Fig. 3*a*. The single, relatively narrow signal with a chemical shift of 61.8 p.p.m.

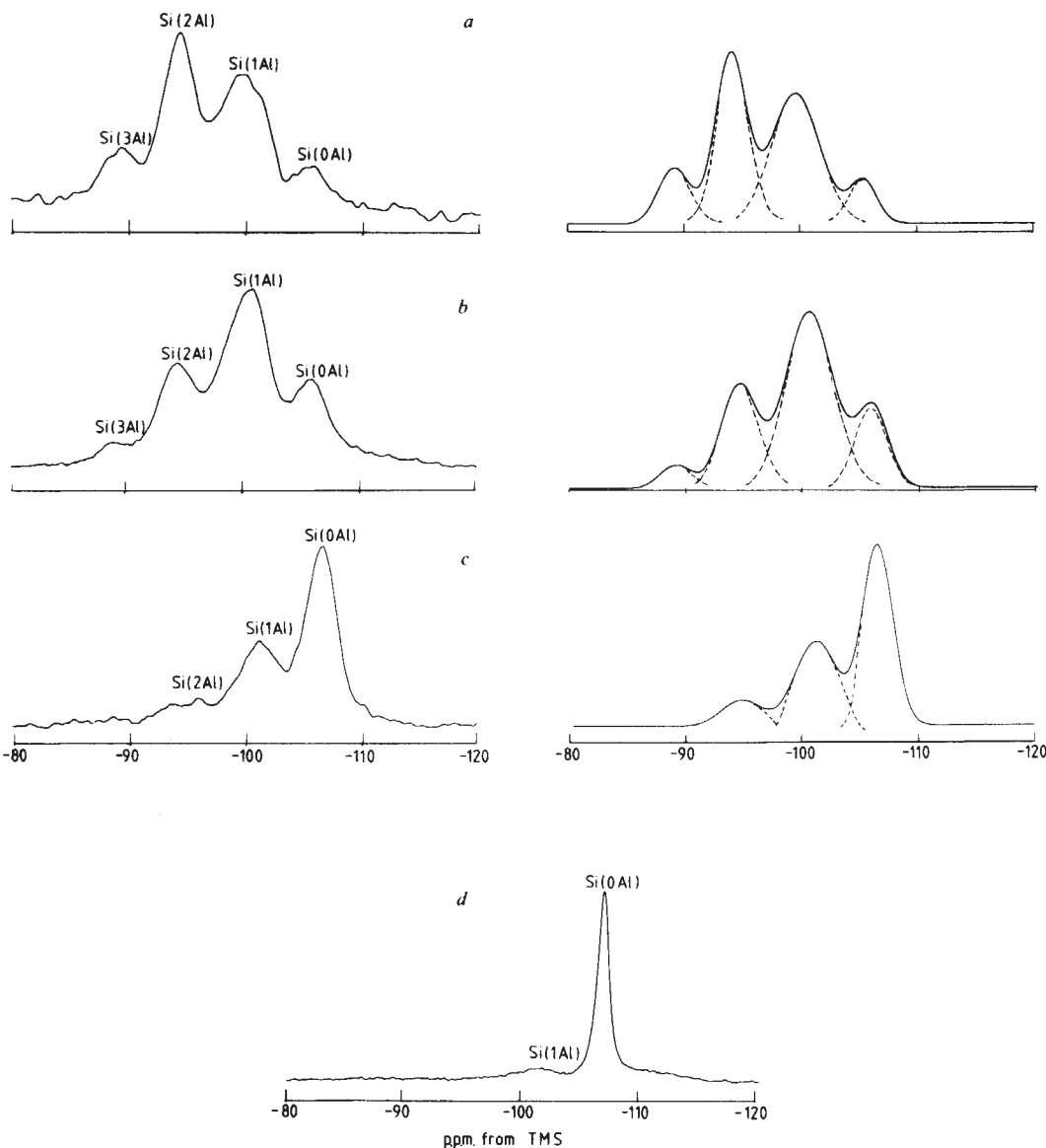


Fig. 2 High-resolution ^{29}Si MASNMR spectra at 79.80 MHz. Experimental spectra are given in the left-hand column; computer-simulated spectra based on gaussian peak profiles and corresponding with each experimental spectrum are given in the right-hand column. Individual deconvoluted peaks are shown as dotted lines. *a*, $\text{NH}_4\text{-Na-Y}$ zeolite (sample 1). *b*, After calcining in air for 1 h at 400 °C (sample 2). *c*, After heating to 700 °C for 1 h in the presence of steam (sample 3). *d*, After repeated ion exchange, heating and prolonged leaching with nitric acid (sample 4).

corresponds to the tetrahedrally bound aluminium (the small humps on the baseline, equidistant from the peak are spinning sidebands arising from chemical shift anisotropy).

Sample 2 was prepared by calcining sample 1 in air for 1 h at 400 °C. Its ^{29}Si spectrum (Fig. 2*b*) is significantly different from that for sample 1, showing a considerable decrease in the intensity of Si(3Al) and Si(2Al) signals with respect to Si(0Al). The calculated intensity ratio is now 4.8:24.6:55.3:15.3, leading to Si/Al ratio of 3.37 in the framework. Chemical analysis shows, however, no change in composition; clearly MASNMR is a superior analytical tool for the study of the status and quantity of chemical elements in the solid state. The observed increase in Si/Al ratio signifies that 23.3% of aluminium has been removed from the framework, with a corresponding loss of ion-exchange capacity. The ^{27}Al spectrum (Fig. 3*b*) is also markedly changed, showing two distinct signals: a large one (at 61.0 p.p.m.) corresponding to aluminium still remaining in tetrahedral sites, and a smaller one (at 0.59 p.p.m.) corresponding unmistakably to Al in octahedral coordination. This aluminium is now in the interstitial cationic positions, the ratio, r , of the intensities of the tetrahedral and the octahedral signals being 5.44 (the curve-fitted peaks were found to be exactly lorentzian). The total amount of aluminium in sample 2 is the same as in sample 1, as no washing was involved. Simple calculation leads to the expression $\text{Si}/\text{Al} = 2.61(r+1)/r = 3.09$ for the composition of the tetrahedral framework. This value

is lower than 3.37 calculated from the ^{29}Si spectrum, but the difference corresponds to only three Al atoms in a unit cell containing 192 tetrahedral atoms. The intensities of the signals in both the ^{29}Si and ^{27}Al spectra do not display appreciable change over a 10-fold increase in the delay time of the experiment, which indicates that they are quantitatively reliable.

Sample 3 was prepared by heating $\text{NH}_4\text{-Na-Y}$ to 700 °C for 1 h in the presence of steam. The ^{29}Si spectrum (Fig. 2*c*) is completely transformed with respect to that for samples 1 and 2. The Si(3Al) peak has disappeared, and the intensity ratio of the Si(2Al), Si(1Al) and Si(0Al) signals is now 10.7:36.7:52.6, leading to the framework composition $(\text{Si}/\text{Al})_{\text{NMR}} = 6.89$. The ^{27}Al spectrum (Fig. 3*c*) shows considerable amounts of octahedrally coordinated aluminium, as evidenced by the resonance at 0.0 p.p.m. There is also some amorphous material responsible for the raised baseline; it is therefore difficult to calculate the ratio of intensities reliably. We estimate $r = 1.62$, giving $\text{Si}/\text{Al} = 4.20$, lower than the value obtained from the ^{29}Si spectrum. The interstitial Al could be only partly octahedral; indeed, by assuming that one-third of it is tetrahedral, in accordance with the reported spinel structure for γ -alumina^{11,20}, we obtain agreement between the framework Si/Al ratios calculated from ^{29}Si and ^{27}Al MASNMR, respectively. The same holds for the small difference observed for sample 2.

Sample 4 was prepared by repeated application of the procedure used for preparing sample 3, followed by prolonged

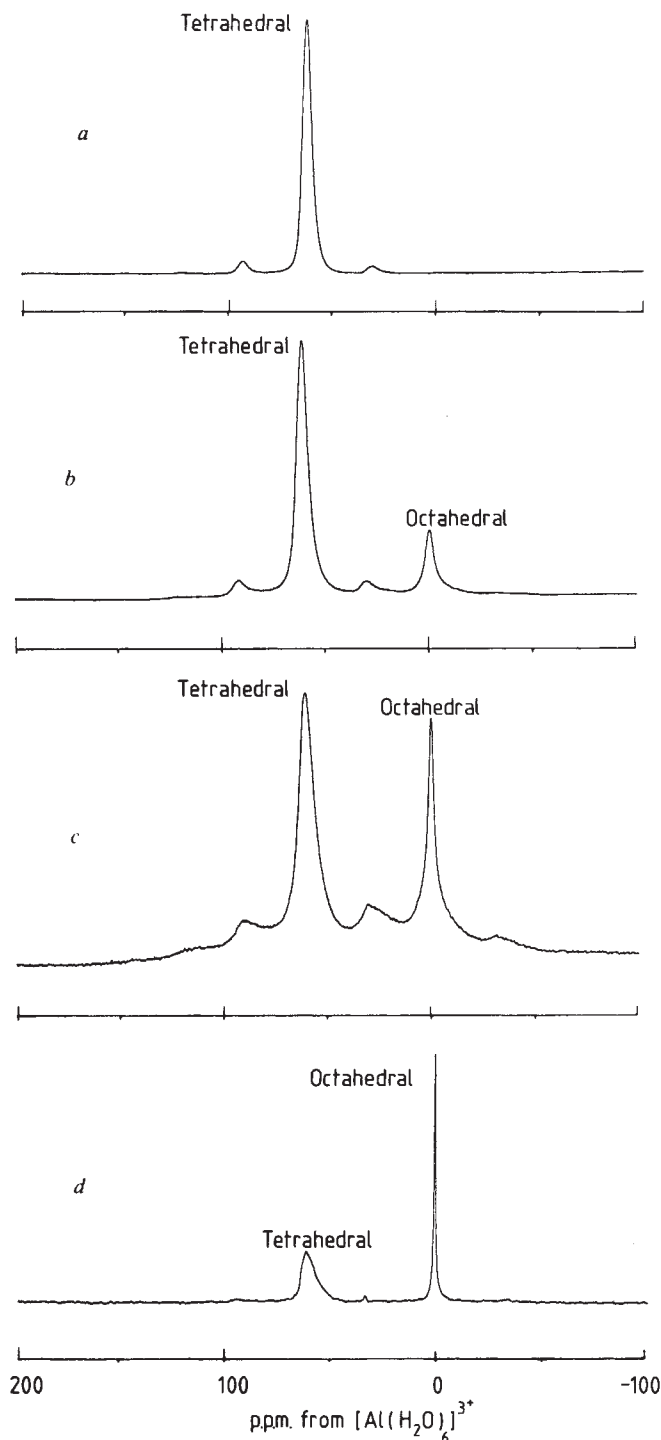


Fig. 3 ^{27}Al MASNMR spectra at 104.22 MHz. *a*, Sample 1. Small humps equidistant from the peak are spinning sidebands. *b*, Sample 2. *c*, Sample 3. *d*, Sample 4.

leaching with 1 M nitric acid. The product has very low aluminium content, with $\text{Si}/\text{Al} > 50$ (by chemical analysis) and a unit cell parameter $a = 24.31 \text{ \AA}$. The ^{29}Si spectrum (Fig. 2d) shows one sharp peak (FWHM = 1.1 p.p.m.) at -106.9 p.p.m. and a very small broad signal at $\sim -101.3 \text{ p.p.m.}$, clearly attributable to some residual Si(1Al) units. The sample contains amorphous material (note the slight elevation of the baseline), but its bulk is evidently very crystalline, the conclusion supported by the very sharp X-ray diffractogram and IR spectrum. The ^{29}Si MASNMR spectrum strongly suggests that framework vacancies created by the removal of aluminium are subsequently re-occupied by silicon. If this were not the case, the spectrum

would be more complex, reflecting a range of possible environments for Si atoms including one, two or three neighbouring hydroxyl groups. Cross-polarization, involving $\text{Si}-\text{O}-\text{H}$ bonds, would prove illuminating here. We note that the chemical shift of the Si(4Si) signal in sample 4 is very close to that reported for quartz⁵ (-107.4 p.p.m.).

The ^{27}Al MASNMR spectrum of sample 4 (Fig. 3d) contains a broad (FWHM = 7.1 p.p.m.) tetrahedral peak at 61.0 p.p.m. and an extremely sharp (FWHM = 0.8 p.p.m.) octahedral signal at 0.23 p.p.m. The latter, the sharpest ^{27}Al signal we have measured in any solid, is due to motionally free $[\text{Al}(\text{H}_2\text{O})_6]^{3+}$ in the cationic positions, not removed by leaching with acid. The ratio of intensities of the two signals is ~ 3 . Such octahedral aluminium that still remains, neutralizes the residual negative framework charge caused by the presence of tetrahedrally coordinated aluminium. If all $[\text{Al}(\text{H}_2\text{O})_6]^{3+}$ ions are exchangeable, an ion-exchange equilibrium is established during acid leaching between $[\text{Al}(\text{H}_2\text{O})_6]^{3+}$ and H_3O^+ competing for the cationic sites. Such equilibrium generally favours multivalent cations. (Note that one octahedral Al atom balances the charge of three tetrahedral Al atoms.)

Discussion

A recent study involving sorption measurements of gases²¹ showed that materials prepared in a manner similar to that for sample 4 contain a secondary mesopore system with pore radii in the range 15–19 Å, suggesting that the silicon which re-occupies empty tetrahedral sites does not come exclusively from the surface or from amorphous parts of the crystals, but also from their bulk, probably involving the elimination of the entire sodalite cages. MASNMR seems to support this conclusion. The Si(4Si) peak in Fig. 2d, although sharp, is almost twice the width of the analogous peak that we measured^{9,10} in the highly siliceous faujasite prepared by treating zeolite Y with gaseous SiCl_4 , and of very similar aluminium content. This suggests that the structure of the thermally dealuminated material is less perfect than that of faujasite dealuminated using SiCl_4 .

We acknowledge support from the Universities of Cambridge and Guelph and from BP Research Centre, Sunbury. C.A.F. acknowledges the financial support of the Natural Sciences and Engineering Research Council (NSERC) of Canada in the form of an operating grant, G.C.G. the award of a NSERC graduate scholarship. The NMR spectra were obtained at the South-West Ontario High Field NMR Facility (Dr R. E. Lenkinski, Manager) funded by a Major Installation grant from NSERC.

Received 29 December 1981; accepted 2 February 1982.

1. Milton, R. M. U.S. Patent 2,882,244 (1959).
2. Breck, D. W. U.S. Patent 3,130,007 (1964).
3. McDaniel, C. V. & Maher, P. K. *Molecular Sieves* (ed. Barrer, R. M.) 186–195 (Society of Chemical Industry, London, 1968).
4. McDaniel, C. V. & Maher, P. K. U.S. Patent 3,449,070 (1969).
5. Lippmaa, E., Mägi, M., Samoson, A., Engelhardt, G. & Grimmer, A.-R. *J. Am. chem. Soc.* **102**, 4889–4893 (1980).
6. Ramdas, S., Thomas, J. M., Klinowski, J., Fyfe, C. A. & Hartman, J. S. *Nature* **292**, 228–230 (1981).
7. Klinowski, J., Thomas, J. M., Fyfe, C. A. & Hartman, J. S. *J. phys. Chem.* **85**, 2590–2594 (1981).
8. Lippmaa, E., Mägi, M., Samoson, A., Tarmak, M. & Engelhardt, G. *J. Am. chem. Soc.* **103**, 4992–4996 (1981).
9. Klinowski, J. *et al. JCS Chem. Commun.* 570–571 (1981).
10. Klinowski, J., Thomas, J. M., Fyfe, C. A., Gobbi, G. C. & Hartman, J. S. *Inorg. Chem.* (submitted).
11. Klinowski, J., Ramdas, S., Thomas, J. M., Fyfe, C. A. & Hartman, J. S. *JCS Faraday Trans. II* (in the press).
12. Engelhardt, G., Lohse, U., Lippmaa, E., Tarmak, M. & Mägi, M. *Z. anorg. allg. chem.* **482**, 49–64 (1981).
13. Müller, D., Gessner, W., Behrens, H.-J. & Scheler, G. *Chem. phys. Lett.* **79**, 59–62 (1981).
14. Fyfe, C. A., Gobbi, G. C., Hartman, J. S., Klinowski, J. & Thomas, J. M. *J. phys. Chem.* (in the press).
15. Kerr, G. T. *Molecular Sieves* (eds Meier, W. M. & Uytterhoeven, J. B.) 219–229 (Advances in Chemistry Series No. 121, American Chemical Society, 1974).
16. Galliezot, P., Beaumont, R. & Barthomeuf, D. *J. phys. Chem.* **78**, 1550–1553 (1974).
17. Breck, D. W. & Skeels, G. W. in *Proc. 5th int. Conf. on Zeolites Naples* (ed. Rees, L. V. C.) 335–343 (Hayden, London, 1980).
18. Fyfe, C. A. *et al. J. Magn. Res.* (submitted).
19. Andrew, E. R. *Int. Rev. phys. Chem.* **1**, 195–224 (1981).
20. Wells, A. F. *Structural Inorganic Chemistry* (Oxford University Press, 1962).
21. Lohse, U. *et al. Z. anorg. allg. Chem.* **460**, 179–190 (1980).

Construction of a functional human suppressor tRNA gene: an approach to gene therapy for β -thalassaemia

Gary F. Temple^{*‡}, Andrée M. Dozy^{*}, Kenneth L. Roy[†] & Yuet Wai Kan^{*}

^{*} Howard Hughes Medical Institute Laboratory, and Departments of Medicine, Biochemistry and Biophysics, University of California, San Francisco, California 94143, USA

[†] Department of Microbiology, University of Alberta, Edmonton, Alberta, Canada T6G 2E9

A human tRNA^{Lys} gene was converted to an amber suppressor by site-specific mutagenesis of the anticodon. The mutated tRNA^{Lys} gene directed synthesis of a tRNA that suppressed the UAG amber nonsense mutation in β^0 thalassaemia mRNA. Such genes may be used to detect other nonsense mutations in mammalian cells and may provide an approach to gene therapy for β^0 thalassaemia due to nonsense mutations.

THE thalassaemias are a heterogeneous group of hereditary disorders characterized by defective globin chain synthesis¹. They are broadly classified into two major subgroups, α and β thalassaemia, depending on whether α or β chain synthesis is impaired. Molecular studies reveal that many different kinds of mutations produce the thalassaemia phenotype. These lesions include structural gene deletions¹, nucleotide substitutions in the intervening sequences that cause abnormal mRNA processing^{2,3}, and nonsense mutations in the coding region that lead to premature termination of globin chain production^{4,5}.

We previously characterized two different mutations that cause β^0 thalassaemia, a form of thalassaemia in which no β -globin chains are synthesized. In a Chinese patient, a AAG to UAG mutation converts the β^{17Lys} codon to an amber termination codon⁴, and in a Sardinian, a CAG to UAG mutation converts β^{39Gln} to a terminator⁵. The latter mutation has been found in β^0 thalassaemia from other parts of Italy, Algeria and Morocco^{6,7}. The nonsense mutation in β^0 thalassaemia is suppressible in cell-free translation systems with a yeast amber-suppressor tRNA *in vitro*^{5,8}. If a functional suppressor tRNA gene were available, one could test its ability to overcome the nonsense mutation in β thalassaemia *in vivo*. We have now tested the feasibility of this approach by constructing a human suppressor tRNA gene and assessing its ability to suppress the nonsense mutation in β^0 thalassaemia *in vivo* in an oocyte system.

Transcription of human tRNA genes

A 1-kilobase fragment of human DNA that contains a tRNA^{Lys} and a tRNA^{Gln} gene has been isolated and completely sequenced (K. L. Roy, H. Cooke and R. Buckland, in preparation). The sequence of the tRNA^{Lys} gene is identical to the RNA sequence previously reported for rabbit tRNA^{Lys}₃ (ref. 9). Neither gene contains intervening sequences. Based on their DNA sequences, tRNA transcribed from tRNA^{Lys} and tRNA^{Gln} should be 73 and 72 nucleotides long respectively, before the trinucleotide CCA is added to their 3' termini.

The sequence homology between the human tRNA^{Lys} gene and the rabbit tRNA^{Lys}₃ suggests, but does not prove, that the tRNA^{Lys} gene is transcribed and the tRNA is functional. No information is available for the tRNA^{Gln} gene. To test whether these two genes are transcribed *in vivo*, we injected them into the nuclei of *Xenopus laevis*, isolated the RNAs from the oocytes, and analysed the newly synthesized RNA on a polyacrylamide gel (Fig. 1a)^{10,11}. The major RNA products were 76 and 75 nucleotides long respectively, with some longer minor bands also visible. These findings suggested that the tRNA genes were transcribed and that the transcripts were processed.

[‡] Present address: Bethesda Research Laboratories, Gaithersburg, Maryland, USA.

The lengths of the major tRNA bands indicated that CCA trinucleotides were added to the 3' termini of both genes. These bands were eluted from the gel, and their identities confirmed by fingerprinting (data not shown).

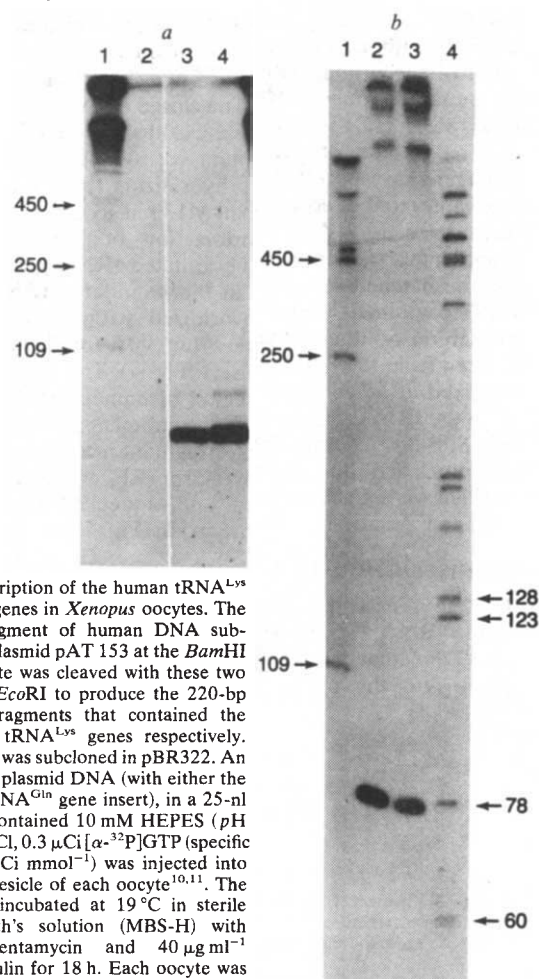


Fig. 1 Transcription of the human tRNA^{Lys} and tRNA^{Gln} genes in *Xenopus* oocytes. The 1-kilobase fragment of human DNA subcloned in the plasmid pAT 153 at the *Bam*HI and *Hind*III site was cleaved with these two enzymes and *Eco*RI to produce the 220-bp and 800-bp fragments that contained the tRNA^{Gln} and tRNA^{Lys} genes respectively. Each fragment was subcloned in pBR322. An aliquot of 7 ng plasmid DNA (with either the tRNA^{Lys} or tRNA^{Gln} gene insert), in a 25- μ l solution that contained 10 mM HEPES (pH 7.9), 60 mM KCl, 0.3 μ Ci [α -³²P]GTP (specific activity 450 mCi mmol⁻¹) was injected into the germinal vesicle of each oocyte^{10,11}. The oocytes were incubated at 19°C in sterile modified Barth's solution (MBS-H) with 75 μ g ml⁻¹ gentamycin and 40 μ g ml⁻¹ bovine γ -globulin for 18 h. Each oocyte was then crushed and lysed in a 30- μ l solution containing 100 mM Tris-HCl (pH 7.5), 300 mM NaCl, 10 mM EDTA, and 1 mg ml⁻¹ Proteinase K, and incubated at 25°C for 45 min. The RNA was twice extracted in phenol, precipitated in ethanol and electrophoresed on a 10% acrylamide gel containing 8 M urea¹². Autoradiography was performed for 3–18 h at -80°C. *Gel a*, (1) DNA size markers (PM2 digested with *Hind*III); and RNA from oocytes injected with (2) pBR322, (3) and (4) pBR322 recombinants with the tRNA^{Gln} and tRNA^{Lys} genes respectively. *Gel b*, injection experiments similar to those shown in *a* performed with the RF of M13mp7 recombinants: (1) DNA size marker (PM2 digested with *Hind*III), (2) tRNA^{Lys} gene, (3) the mutated tRNA^{Lys} gene, (4) DNA size marker (M13mp7 digested with *Msp*I).

Mutation of the anticodon of the tRNA^{Lys} gene

We first mutated the anticodon region of the tRNA^{Lys} gene to test its ability to function as an amber suppressor tRNA gene. The anti-sense strand of the tRNA^{Lys} gene has the sequence AAA at the anticodon. To convert this gene so the tRNA will read the amber terminator UAG, we changed the AAA sequence to TAG using a synthetic pentadecamer (Bio Logicals) which was complementary to the anti-sense strand of the tRNA gene at the anticodon stem and loop, except for the two bases we intended to mutate. The sequences of the anticodon loop (anti-sense strand) of the tRNA gene and the primer used are shown in Fig. 2.

The strategy for mutagenesis is detailed in Fig. 3. The 800-base pair (bp) DNA fragment that contained the tRNA^{Lys} gene was subcloned into M13mp7 phage and single-stranded phage were isolated. The synthetic pentadecamer was used to prime synthesis of the complementary strand with DNA Polymerase I in the presence of T₄ DNA ligase *in vitro*, according to the method of Gillam and Smith^{17,18}. Initial synthesis was at 0 °C to maximize priming despite the mismatch; the temperature was then raised to 25 °C. The double-stranded closed circular forms were isolated on an agarose gel in the presence of ethidium bromide, and used to transfect *Escherichia coli*. Single-stranded phage were isolated from the culture and used as the template for a second round of synthesis which was carried out at 25 °C, without the initial 0 °C incubation, to favour hybridization of the primer with those phage that already contained the mutated sequence^{18,19}. The closed circular double-stranded forms were again isolated and used to transfect *E. coli*.

After a total of five rounds of synthesis, the success of mutagenesis was tested by hybridizing the ³²P-labelled pentadecamer to the recombinant M13 phage. Stringent conditions were used to allow hybridization of the primer to the homologous sequence in the mutated tRNA gene, but not to the mismatched sequence in the unmutated tRNA gene. All the recombinant phage hybridized strongly to the labelled primer on Southern analysis, but 90% of them had 300 bp deleted from the 800-bp insert (Fig. 4). The deletion probably occurred at an A-T rich region of human DNA downstream from the tRNA^{Lys} gene. One recombinant which contained the full 800-bp genomic DNA fragment was cultured in large quantities and the replicative form (RF) of DNA was isolated, purified on a CsCl gradient, and sequenced. It contained the intended mutation at the anticodon (Fig. 5).

Transcription of the mutated tRNA^{Lys} gene

To determine whether the mutation affected the transcription of the tRNA^{Lys} gene, we injected the double-stranded RF of M13 DNA that contained the original and the mutated tRNA^{Lys} genes into the nuclei of *Xenopus* oocyte. Both produced

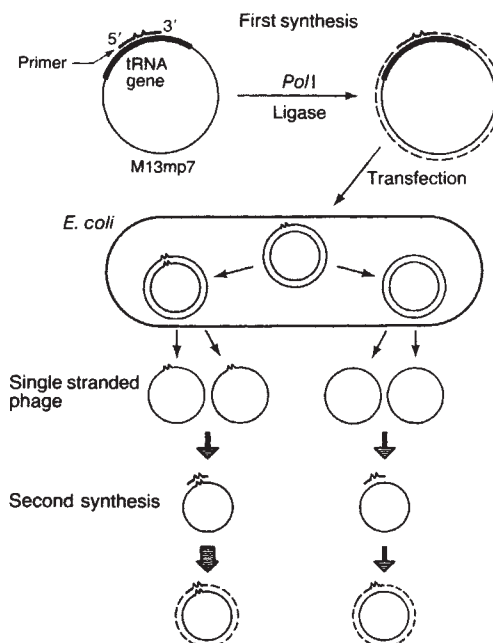


Fig. 3 Strategy of mutagenesis. The 800-bp DNA fragment that contained the tRNA^{Lys} gene with the *Bam*HI and *Eco*RI ends was subcloned into the *Eco*RI site of the bacteriophage M13mp7¹³, after converting the *Bam*HI site to an *Eco*RI site with linkers¹⁴. We selected the recombinant in which the anti-sense strand of the tRNA gene was contained in the single-stranded form of the M13. Single-stranded and RF forms of the recombinant phage were prepared^{15,16}. Site-specific mutagenesis with the pentadecamer was performed according to Gillam and coworkers with modifications¹⁷⁻¹⁹. A 35- μ l mixture that contained 2 pmol of the single-stranded M13 recombinant, 500 pmol of the synthetic pentadecamer, 100 mM NaCl, 40 mM Tris-HCl (pH 7.5), 20 mM MgCl₂, and 2 mM β -mercaptoethanol was heated to 70 °C for 10 min and then cooled slowly over 1 h to 0 °C. A 28- μ l aliquot was removed and the mixture adjusted to a final volume of 70 μ l, which contained 29 mM Tris HCl (pH 7.5), 40 mM NaCl, 14 mM MgCl₂, 1.4 mM β mercaptoethanol, 500 mM each of 4 dNTPs, 240 mM ATP, 90 U ml⁻¹ Klenow fragment of DNA Polymerase I (New England Biolabs) and 100 U ml⁻¹ of T₄ DNA ligase (Bethesda Research Laboratories). The mixture was incubated at 0 °C for 30 min and then at 25-27 °C for 4 h. The DNA was then electrophoresed on a 0.8% agarose gel in the presence of 2 μ g ml⁻¹ ethidium bromide. The band that contained the closed circular double-stranded DNA was excised and eluted with glass beads²⁰. This DNA was used to transform *E. coli*, strain GM48 (*Dcm*⁻, *Dam*⁻), previously made competent with CaCl₂ treatment²¹. 0.3 ml of transformed GM48 cells was mixed with 1 ml of JM103¹³ (at A₆₆₀ of 0.2), and an aliquot was plated to determine the titre and prepare individual plaques for screening. The remainder was grown overnight at 37 °C and single-stranded DNA isolated from the supernatant to serve as template for repeated *in vitro* synthesis. In subsequent rounds of synthesis, the conditions were modified to favour priming of the mutated species by decreasing the primer-to-template ratio to 25 to 1 and omitting the initial incubation at 0 °C^{17,18}. A total of five rounds of synthesis was performed.

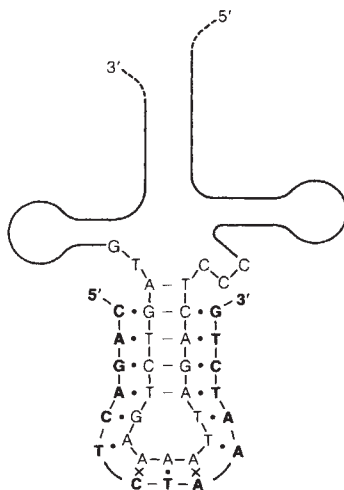
approximately the same amount of 76-bp tRNA products expected for tRNA^{Lys} (Fig. 1b). Thus mutagenesis did not alter transcription of the tRNA^{Lys} gene.

Ability of the mutated tRNA^{Lys} gene to direct suppression of the amber codon

We then tested the ability of the mutated tRNA^{Lys} gene to produce a functional amber suppressor tRNA. We used the mRNA from a patient with β^0 thalassaemia due to a $\beta^{17\text{Lys}}(\text{AAG} \rightarrow \text{UAG})$ mutation. The β/α globin mRNA ratio in this patient's reticulocytes is 0.15 (ref. 4). We previously showed that suppression with a yeast serine-inserting amber suppressor tRNA in a cell-free system produces a β -globin chain with an altered chromatographic elution pattern because the $\beta^{17\text{Lys}}$ is replaced with serine⁸. If the mutated human tRNA^{Lys} gene produced a functional amber suppressor tRNA and inserted lysine at the β^{17} position, we would expect that a β -globin chain with a normal charge would be produced. Also, the length of the β -globin chain would not be affected, since the normal termination codon of β -globin mRNA is UAA²⁵.

Xenopus oocyte nuclei were first injected either with the normal or the mutated tRNA^{Lys} gene. The cytoplasm of the

Fig. 2 Sequences of the anti-sense strand of the tRNA gene at the anticodon stem and loop, and the synthetic pentadecamer with the two-base pair mismatch (X) at the anticodon.



oocytes was injected 20 hours later with polyA⁺ reticulocyte RNA from the β^0 -thalassaemia patient. The oocytes were incubated for another 48 hours in the presence of ³H-histidine and the protein was extracted. The globin chains were immunoprecipitated using a monoclonal antibody, eluted, and analysed on a Triton X-urea gel²⁶. As expected, no β -globin synthesis was detected in the oocytes injected only with the mRNA, or with mRNA plus the normal tRNA^{Lys} gene. However, in the oocytes injected with the mutated tRNA^{Lys} gene, the β^0 thalassaemia mRNA now directed synthesis of an authentic β -globin chain, as identified by immunological reactivity and electrophoretic mobility (Fig. 6). Hence the tRNA^{Lys} gene with the anticodon mutation functioned as an amber suppressor and most likely inserted Lys into the β^{17} position of the β -globin chain.

Discussion

We previously showed that in the two known types of β^0 thalassaemia due to nonsense mutations, the UAG codon can be suppressed *in vitro* by adding a yeast suppressor tRNA^{5,8}. We undertook the present study to see whether this finding could provide an approach for *in vivo* gene therapy. To conduct *in vivo* DNA transformation experiments, we needed a suppressor tRNA gene. Although suppressor tRNA genes have been derived from bacteria and yeast, evidence suggests that some yeast tRNA genes are not processed correctly in mammalian cells (P. Berg, personal communication). Mammalian tRNA genes were therefore preferable for our experiment. Suppressor tRNAs have been identified in rabbit reticulocytes²⁷ and bovine liver²⁸, but since no mammalian suppressor tRNA genes have yet been isolated, we constructed one *in vitro*. We mutated the anticodon of a tRNA^{Lys} gene (AAA) to TAG (anti-sense strand) with a primer mismatched at the two-base sequence. We showed that this gene was transcribed and that the transcript functioned as an amber suppressor *in vivo* in *Xenopus* oocyte, because it read through the UAG codon in β^0 -thalassaemia mRNA. It is thus possible to convert a tRNA^{Lys} to an amber suppressor tRNA by mutating the anticodon.

The availability of mammalian suppressor tRNA genes offers a means of detecting nonsense mutations in higher organisms. At present, apart from viruses, bacteria and yeast, the only naturally occurring nonsense mutations known to cause diseases in higher organisms are the two defined in β -thalassaemia in humans⁴⁻⁶. Other nonsense mutations undoubtedly occur; just in the coding region of the β -globin gene there are 29 positions at which a single nucleotide substitution could produce a termination codon. Certain mouse and Chinese hamster cell lines with hypoxanthine guanine phosphoribosyl transferase

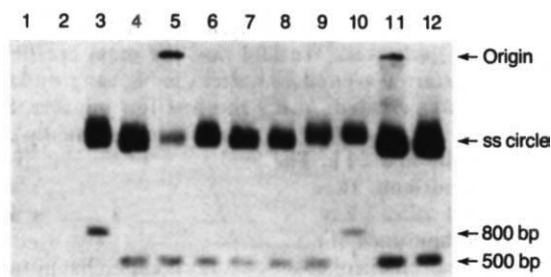


Fig. 4 Southern analysis of M13 recombinant phage. After the fifth round of synthesis, RF DNA from the cell pellet¹⁶ was prepared from individual plaques, and the RF DNA from 35 of the mutated clones was digested with *EcoRI*, electrophoresed on an 0.8% agarose gel and transferred to nitrocellulose filters²². The filters were prehybridized for 2 h, and then hybridized for 16 h with the synthetic pentadecamer, which was previously labelled with ³²P at the 5' end, in a buffer containing 6×SSC, 10× Denhardt's solution at 41 °C and 20 µg ml⁻¹ yeast RNA. The filter was washed three times with 6×SSC for 20 min at 12 °C²³, dried and autoradiographed. Lane 1, DNA from M13; lane 2, M13 containing the unmutated tRNA^{Lys} gene; lanes 3–12, ten of the DNA samples containing the M13 tRNA^{Lys} gene after mutagenesis. All 35 clones isolated hybridized strongly to the probe. In all but two of the clones shown, deletions shortened the insert from 800 to 500 bp. The DNA from the clone shown in lane 10 was grown in large volume for further studies.

Fig. 5 Sequencing gel showing the mutated sequence in the anticodon region of the tRNA^{Lys} gene. The RF DNA prepared from the mutated clone shown in Fig. 4 was purified by CsCl banding for 48 h. A *KpnI* site, 55 bp from the anticodon was labelled with ³²P at the 3' terminus using T₄ DNA polymerase²⁴, and the sequence of the tRNA gene was determined by the method of Maxam and Gilbert on a 15% polyacrylamide gel in 8 M Urea¹².

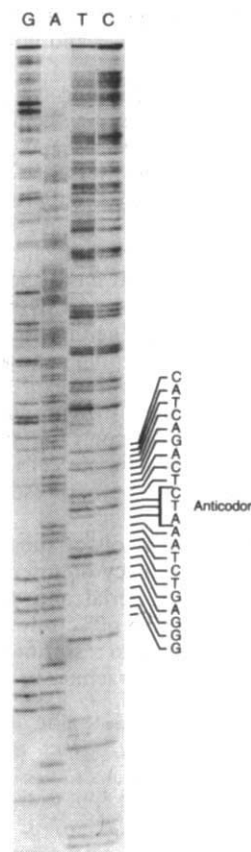
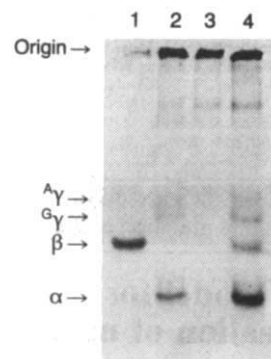


Fig. 6 Analysis of globin chain synthesis in *Xenopus* oocyte after injection with tRNA genes and polyA⁺ mRNA from human reticulocytes and immunoprecipitation of the globin with monoclonal antibodies. A sample of 7 ng DNA that contained either the normal or mutated tRNA^{Lys} gene was injected into the nucleus of each *Xenopus* oocyte as described in Fig. 1, but without [α -³²P]GTP. 20 h later, 10 ng of polyA⁺ reticulocyte RNA in 50 nl of 0.01 M Tris-HCl (pH 7.5), 0.02 M KCl, and 0.25 mM haemin was injected into the cytoplasm at the equatorial band and incubation was continued at 19 °C for 48 h after adding 0.2 mCi ³H-histidine to each ml of the medium. Oocytes were washed twice in the incubation buffer, then squashed in a 5-µl solution per oocyte of 100 mM Tris-HCl (pH 6.8), 150 µg ml⁻¹ phenyl methyl sulphonyl fluoride (PMSF), and 2.5 mM EDTA. The lysate was clarified twice by centrifugation at 10,000 r.p.m. in an Eppendorf centrifuge. The extract from three oocytes was pooled and reacted with 25 µg of a monoclonal antibody against human globin in a 50 µl TENT buffer (25 mM Tris [pH 7.4], 10 mM EDTA, 350 mM NaCl, 0.15% Triton X-100, 1 mg ml⁻¹ bovine serum albumin). (This antibody reacts most strongly with human β -chains, but also reacts with α - and γ -globin chains.) After 18 h at 4 °C, 50 µl of a 10% suspension of *Staphylococcus A* in TENT buffer was added, and incubated at 4 °C for 6 h. The *Staphylococcus A* was pelleted by centrifugation at 10,000 r.p.m. for 2 min at 4 °C, and washed three times with the TENT buffer (without bovine serum albumin). Haemoglobin A 1 µg was added as carrier and the bound radioactive globin was eluted by heating the pellet to 90 °C for 2 min in a 25-µl buffer containing 6 M urea (BioRad, electrophoretic purity), 1 M 2-mercaptoethanol (Sigma), 2% Triton X-100, 0.33 mg ml⁻¹ Pyronin-Y. The cells were pelleted and the supernatant applied to a Triton X urea gel²⁶. The gels were dried, soaked in Enhance (New England Nuclear) and autoradiographed for 4 days. Lane 1, oocytes injected with normal β -globin mRNA; lane 2, injected with β^0 mRNA alone; lane 3, injected with the mutated tRNA^{Lys} gene only; lane 4, injected with β^0 mRNA and the mutated tRNA^{Lys} gene. When the β^0 mRNA was injected together with the original tRNA^{Lys} gene, the result was identical to that shown in lane 2. The migration of the α -, β -, γ -, and δ -globin chains was determined with unlabelled globin markers.



(HGPRT) deficiency are also suspected to be due to nonsense mutations^{29,30}, and nonsense mutations have been detected in *Drosophila*³¹ and *Caenorhabditis elegans*^{32,33} by genetic analysis. Transforming cells from such organisms with suppressor tRNA genes may confirm the presence of nonsense mutations in these species and detect such mutations in other human genetic disorders.

Insertion of suppressor tRNA genes can be tested as an approach to gene therapy in β -thalassaemia. While this approach would be limited to β -thalassaemias due to nonsense mutations, it has a theoretical advantage over other proposed gene therapies for β -thalassaemia, such as insertion of a normal β -globin gene into thalassaemia cells. In experiments where a cloned β -globin gene was inserted into heterologous cells, the level of β -globin expression was low³⁴. At present, the mechanism that controls normal β -globin gene expression is not understood. In contrast, since tRNAs are essential for protein synthesis in all cells, inserted functional tRNA genes are likely to be expressed, even in differentiated cells such as erythroid cells.

The success of suppressor tRNA gene therapy depends on the efficiency of suppression by the mutated tRNA. If, as in some bacterial and yeast suppressor tRNAs, the efficiency approaches 60% (refs 35, 36), then the gene should significantly contribute to restoring the balance of α - and β -globin synthesis in β -thalassaemia. In addition, the efficiency of suppression of

suppressor tRNAs can vary. It is therefore important to test the effects of multiplicity of gene insertion and the degree of suppression with various tRNAs. We are currently mutating the tRNA^{Gln} gene in a similar manner in order to study the Mediterranean type of β -thalassaemia due to a $\beta^{39\text{Gln(CAG)}\rightarrow\text{UAG}}$ mutation⁵.

Of course, before the use of suppressor tRNAs can be contemplated for gene therapy, their effect on other cell functions must be assessed. An amber-suppressor tRNA could also suppress termination of proteins that utilize the UAG codon as their normal terminator. Although suppressor tRNAs are lethal in certain yeast strains, no information is available on their effects on cells of higher organisms. The availability of a functional suppressor tRNA gene allows us to test for any deleterious effects it may have on mammalian cells.

We thank Dr Rae Lyn Burke for suggesting the use of an ethidium bromide agarose gel to separate the different forms of synthetic M13 products, Drs Michael Smith and U. L. RajBhandary for their advice on mutagenesis, Dr George Stamatoyannopoulos for the antibody against human globin, and Dr Joe Hedgepeth for the GM48 strain of *E. coli* and J. Gampell for editorial comments. This research was supported by grants from NIH (AM 16666), the March of Dimes/Birth Defects Foundation, and the Natural Sciences and Engineering Research Council of Canada. Y.W.K. is an investigator of the Howard Hughes Medical Institute.

Received 15 September 1981; accepted 23 February 1982

- Weatherall, D. J. & Clegg, J. B. *The Thalassaemia Syndromes* (Blackwell, Oxford, 1981).
- Spritz, R. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 2455–2459 (1981).
- Westaway, D. & Williamson, R. *Nucleic Acids Res.* **9**, 1777–1788 (1981).
- Chang, J. C. & Kan, Y. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2886–2889 (1979).
- Trecartin, R. F. *et al. J. clin. Invest.* **68**, 1012–1017 (1981).
- Pergolizzi, R. *et al. Nucleic Acids Res.* **9**, 7065–7072 (1981).
- Orkin, S. H. & Goff, S. C. *J. biol. Chem.* **256**, 9782–9784 (1981).
- Chang, J. C., Temple, G. F., Trecartin, R. F. & Kan, Y. W. *Nature* **281**, 602–603 (1979).
- Raba, M. *et al. Eur. J. Biochem.* **97**, 305–318 (1979).
- Gurdon, J. B., Lane, C. D., Woodland, H. R. & Marbaix, G. *Nature* **233**, 177–182 (1971).
- Gurdon, J. B., Partington, G. A. & De Robertis, E. M. *J. Embryol. exp. Morph.* **36**, 541–553 (1976).
- Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
- Messing, J., Crea, R. & Seeberg, P. H. *Nucleic Acids Res.* **9**, 309–321 (1981).
- Goodman, H. M. *Meth. Enzym.* **65**, 63–64 (1980).
- Winter, G. & Fields, S. *Nucleic Acids Res.* **8**, 1965–1974 (1980).
- Birnboim, H. D. & Doly, J. *Nucleic Acids Res.* **7**, 1513–1523 (1979).
- Gillam, S. & Smith, M. *Gene* **8**, 81–97 (1979).
- Gillam, S. & Smith, M. *Gene* **8**, 99–106 (1979).
- Gillam, S., Astell, C. R. & Smith, M. *Gene* **12**, 129–137 (1980).
- Vogelstein, B. & Gillespie, D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 615–619 (1979).

- Cohen, S. N., Chang, A. C. Y. & Hsu, L. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2110–2114 (1972).
- Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
- Wallace, R. B. *et al. Nucleic Acids Res.* **6**, 3543–3557 (1979).
- Englund, P. *Meth. Enzym.* **65**, 39 (1980).
- Marotta, C. A., Wilson, J. T., Forget, B. G. & Weissman, S. M. *J. biol. Chem.* **252**, 5040–5053 (1977).
- Alter, B. P., Goff, S. C., Efremov, G. D., Gravely, M. E. & Huisman, T. H. J. *Br. J. Haematol.* **44**, 527–534 (1980).
- Geller, A. I. & Rich, A. *Nature* **283**, 41–46 (1980).
- Diamond, A., Dudock, B. & Hatfield, D. *Cell* **25**, 497–506 (1981).
- Capecchi, M. R., Von der Haar, R. A., Capecchi, N. E. & Sveta, M. M. *Cell* **12**, 371–381 (1977).
- Caskey, C. T., Fenwick, R. G., Kruh, G. & Konecki, D. in *Nonsense Mutations and tRNA Suppressors* (eds Celis, J. E. & Smith, J. D.) 235–254 (Academic, New York, 1979).
- Suzuki, D., Kaufman, T. & Falk, D. in *Genetics and Biology of Drosophila* (eds Ashburner, M. & Novitsky, E.) 213–315 (Academic, New York, 1976).
- Waterston, R. H. & Brenner, S. *Nature* **275**, 715–719 (1978).
- Waterston, R. H. *Genetics* **97**, 307–325 (1981).
- Wold, B. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 5684–5688 (1979).
- Capecchi, M. R., Hughes, S. H. & Wahl, G. M. *Cell* **6**, 269–277 (1975).
- Gesteland, R., Wolfner, M., Grisafi, P., Fink, G., Botstein, D. & Roth, J. R. *Cell* **7**, 381–390 (1976).

LETTERS TO NATURE

A model for the cosmic creation of nuclear exergy

K.-E. Eriksson*†, S. Islam* & B.-S. Skagerstam†

* Physical Resource Theory Group and † Institute of Theoretical Physics, Chalmers University of Technology and University of Göteborg, S-412 96 Göteborg, Sweden

The exergy¹ of a physical system is the maximum amount of mechanical work that can be extracted from that system so that a system in thermodynamic equilibrium has zero exergy. An old theme in cosmic thermodynamics is that the second law implies that the Universe is running down (running out of exergy) and approaching thermodynamic equilibrium, 'heat death'. However, in the standard model of the early Universe², at 0.01 s after the big bang the Universe consisted of ordinary matter (nucleons), electrons and positrons, neutrinos, photons and gravitons. The gravitons were decoupled, but the others were in thermodynamic equilibrium. Thus the Universe had already reached the state of heat death, and its exergy was zero. The main problem is, therefore, not to describe the running-down of the Universe, but to understand its revival. When and

how was exergy created, in particular the nuclear exergy, which is transformed into life-supporting light in our Sun? We have studied a model which should represent the nucleon gas of the early Universe quite well. We find that the main creation of nuclear exergy started around 10 s after the big bang, and most of the exergy was created during the first few minutes, 85% during the first hour, and that the process was essentially completed during the first 24 h. The final value of the exergy was 7.72 MeV per nucleon. The consumption of this exergy started much later and takes place in the stars over a time scale of hundreds and thousands of millions of years.

Although the problem of cosmic exergy creation has not been widely discussed, general answers have been given by Tolman³, and later by Layzer⁴ and Landsberg⁵. The solution is that during the expansion of the Universe, the conditions for local equilibrium changed with changing temperature and pressure. The reactions tending towards equilibrium, however, lagged behind, and disequilibrium arose. This does not violate the second law of thermodynamics. If one considers any element $\Delta\Omega$ of the Universe, its entropy ΔS increases, but its maximum attainable entropy ΔS_{max} may increase faster due to the expansion, which thus induces a growing disequilibrium.

Work can be extracted from a system only to the extent that this deviates from equilibrium. Exergy is thus a measure of deviation from equilibrium, of contrast, internally or against

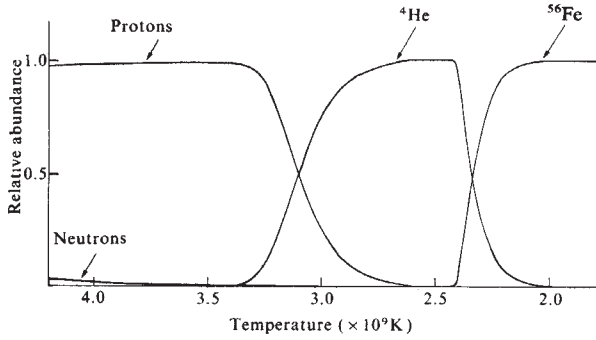


Fig. 1 Relative abundances $P_i^{(0)}$ at equilibrium for the four nucleon states considered in the nucleon gas model.

an external background^{1,6-9}. For instance, in an environment of temperature T_0 , a package of heat ΔQ at temperature T has the exergy

$$\Delta E = \Delta Q \left(1 - \frac{T_0}{T} \right) \quad (1)$$

that is here exergy equals energy times a quality factor, which is simply the Carnot factor $1 - T_0/T$. As the quality may decrease (with decreasing contrast), unlike energy, exergy is not conserved. Physical processes always go towards increasing entropy which usually means that systems move towards equilibrium and thus towards decreasing exergy. Therefore to attain a state away from equilibrium or to sustain such a state, an exergy input is required.

The totally dominant exergy source of the Earth is the inflow of solar radiation, the value of which derives from its contrast with the ambience of the Solar System, the cold interstellar space. The solar radiation derives from a continuous conversion of nuclear exergy in the interior of the Sun. Our problem is to understand how this exergy arose through the mentioned mechanism. In the nucleon gas of the early Universe, exergy and entropy increased together, and we shall study a simple statistical model for this process.

The era we are considering is dominated by radiation, and the photon gas may be considered as a heat bath which is continuously cooling down. The nucleon gas is always in thermal contact with this bath. Below a certain temperature (2.0×10^9 K) an iron nucleus (^{56}Fe)—if it could be formed by a coincidence of events (extremely unlikely) or be introduced from outside—would be thermally stable in contact with the heat bath. From then on the nucleon gas may be considered as being overheated, the natural phase at this temperature being a gas of iron nuclei. (The overheating started much earlier, first through an overabundance of neutrons, which decay only slowly. Also for a short period at temperatures near 2.5×10^9 K (well before helium synthesis) ^4He defines the natural phase of the nucleon gas.)

Several bottlenecks prevent the formation of iron nuclei, the first being the low binding energy of the deuteron. After ^4He has been formed (through ^2H , ^3H and ^3He) at a temperature below the dissociation temperature of deuterons, density and temperature are already below the values required to sustain continued nuclear fusion reactions beyond helium.

The exergy is thus created, when the nucleon gas makes the transition from the equilibrium state to the overheated state. Some exergy is consumed in neutron decay and in helium formation. In the latter practically all remaining neutrons are absorbed. Helium formation is a relatively rapid process, and in our model it occurs at a sharp instant of time.

In the era under consideration the temperature is still high and the density is sufficiently low to consider the nucleon gas as a non-relativistic Maxwell-Boltzmann gas. Consider again an element $\Delta\Omega$ of the gas. The exergy^{10,11} per nucleon is

$$E = kT \sum_i P_i \ln \frac{P_i}{P_i^{(0)}} \quad (2)$$

where i labels states and where P_i is the actual probability

distribution for one nucleon, and $P_i^{(0)}$ is the probability distribution for one nucleon at thermodynamic equilibrium.

We shall consider a simplified model. As helium forms early in the history of the Universe, and because iron is an end product of nuclear burning, we shall include the ground states of ^4He and ^{56}Fe in our model but no other composite nuclei. Thus a nucleon may either be a free proton or a free neutron or it may be bound in ^4He or ^{56}Fe . We shall label these possibilities by an index i of value 1–4.

At low temperatures the statistical factors depend sensitively on the energy levels. This tends to lead to an all-or-none situation. Below the temperature 2×10^9 K, ^{56}Fe totally fixes the normalization of the $P_i^{(0)}$, and inclusion of other nuclides with tighter binding than ^4He would not change the general picture. Thus it is sufficient to include in equation (2) only ^{56}Fe and states that are actually occupied. From those latter states we have excluded the transitional nuclides, ^2H , ^3H , ^3He , which are important for the detailed dynamics, but which, due to their relatively low abundance, would change E as a function of temperature only marginally.

At temperature T the equilibrium distribution of the nucleons over the four nuclear states is

$$P_i^{(0)} = \kappa r \left(\frac{mc^2}{kT} \right)^{3/2} g_i A_i^{5/2} \exp [-A_i(\mu(T) - b_i)/kT] \quad (3)$$

where g_i is the spin multiplicity of states, A_i is the mass number of the nuclide, and b_i is the average binding energy per nucleon (relative to a free proton); m is the nucleon mass, r is the ratio of photon to nucleon abundance in the Universe, κ is a constant, and $\mu(T)$ is the chemical potential.

Table 1 gives A_i , g_i and b_i and further

$$\left\{ \begin{array}{l} m = 938.8 \text{ MeV}/c^2 \\ \kappa = \frac{\sqrt{2\pi}}{8\zeta(3)} = 0.351 \end{array} \right. \quad (4)$$

$$\left\{ \begin{array}{l} \kappa = \frac{\sqrt{2\pi}}{8\zeta(3)} = 0.351 \end{array} \right. \quad (5)$$

The quantity r is not very well known¹²⁻¹⁴. We shall use the value

$$r = 10^9 \quad (6)$$

The results of our calculation do not depend very strongly on the exact value chosen for this parameter.

The chemical potential $\mu(T)$ in equation (3) is fixed by the overall normalization condition

$$\sum_i P_i^{(0)} = 1 \quad (7)$$

The derivation of equation (3) follows standard methods of statistical mechanics.

The actual probability distribution is given in Table 2. It changes abruptly when helium is formed, which in our model takes place instantaneously at the temperature

$$T_0 = 1.0 \times 10^9 \text{ K} \quad (8)$$

The picture would change only marginally if T_0 were changed, and the asymptotic exergy value (7.72 MeV per nucleon) depends on T_0 only through $\gamma(T_0) = \gamma_0$, where $2\gamma_0$ is the fraction of primordial helium, for which we use the value 27% (ref. 2).

In general the neutron fraction is a complicated function, which has to be computed numerically¹⁵⁻¹⁷. The same is true of the relation between temperature and time.

Table 1 Characteristics of the four nuclides considered

i	$A_i Z_i$	g_i	b_i (MeV)
1	^1H	2	0
2	^1_0n	2	-1.30
3	^4He	1	6.43
4	^{56}Fe	1	8.10

Table 2 Actual distribution of nucleons over the four nuclides before and after helium formation

i	$P_i(T > T_0)$	$P_i(T < T_0)$
1	$1 - \gamma(T)$	$1 - 2\gamma(T_0)$
2	$\gamma(T)$	0
3	0	$2\gamma(T_0)$
4	0	0

$\gamma(T)$ is the neutron fraction.

The exergy per nucleon from equation (2) is now obtained from equation (3) and Table 2,

$$E(T) = \mu(T) + |b_2| \gamma(T) - kT \ln \frac{2\kappa r \left(\frac{mc^2}{kT}\right)^{3/2}}{(\gamma(T))^{\gamma(T)}(1-\gamma(T))^{1-\gamma(T)}}; \quad T > T_0 \quad (9)$$

$$E(T) = (1 + 6\gamma_0)\mu(T) - 8\gamma_0 b_3 - kT \ln \frac{2^{8\gamma_0-1} \kappa r \left(\frac{mc^2}{kT}\right)^{3/2}}{\gamma_0^{\gamma_0}(1-\gamma_0)^{1-\gamma_0}}; \quad T < T_0 \quad (10)$$

The equilibrium probability distribution (3) is plotted in Fig. 1. The nucleons dominate completely for $T > 3.5 \times 10^9$ K, and iron dominates for $T < 2 \times 10^9$ K. Helium dominates in an interval around 2.5×10^9 K.

The creation of exergy is shown in Fig. 2. The exergy loss in helium synthesis is ~ 0.6 MeV. Part of this is compensated for in the following evolution, and the net loss is ~ 0.4 MeV. The final value is given by equation (10) in the low temperature limit,

$$E(0) = (1 + 6\gamma_0)b_4 - 8\gamma_0 b_3 = 7.72 \text{ MeV} \quad (11)$$

The nuclear burning in stars with iron as an end product takes place relatively late, and several barriers have to be overcome¹⁸. Only a limited fraction of the nucleons in a star are involved in these reactions. Due to the cold environment the relative exergy loss in the transition from nuclear to radiation exergy in a star is small, of the order of 1% or fractions thereof.

At one stage in the expansion of the Universe matter became so disperse that the thermal contact with surrounding matter

and radiation could no longer prevent density fluctuations from being enhanced by gravitational attraction. Then a new feature entered the picture. When this happened, matter already dominated over radiation in the Universe. The enhancement of density fluctuations meant that the Universe split on several levels into 'compartments' (protogalaxies and protostars) with relatively little interaction with the exterior. Within each compartment the natural reference state is the equilibrium end state for that compartment. Thus at the time of compartmentalization new exergy was created. For a protostar which is to end up as a white dwarf the nuclear exergy from the first few minutes is augmented by some gravitational exergy (of the order of a few keV per nucleon).

For a more massive protostar which is to end up as a cold neutron star there is also a substantial addition of combined nuclear-gravitational exergy of the order of 40 MeV per nucleon¹⁹. The gravitational exergy here gives the dominant contribution. The heavy elements (heavier than iron) formed in a supernova explosion with a neutron star as the final state for its core, thus have exergy which is mainly of this origin.

For a system large enough to have a black hole as its final state, the exergy per nucleon may be a large fraction of its rest energy²⁰.

If matter in the form of nucleons is unstable, as modern elementary particle physics suggests²¹, then matter itself represents a special form of exergy, which was created during the first few microseconds, when an excess of matter over antimatter was developed followed by annihilation of all antimatter and most of the matter²²⁻³¹. Matter as an exergy form is extremely longlived but if the Universe is open and expands for infinity then this exergy will also finally be consumed in nucleon decay (with a half life of $\sim 10^{31}$ yr) and at the end the Universe will consist of ever cooler electrons and positrons (too disperse to annihilate completely), neutrinos, photons and gravitons.

We thank Göran Niklasson and Ture Eriksson for clarifying discussions on the properties of the nuclear gas and on relativistic cosmology. B.-S.S. also thanks J. Prentki and the CERN Theory Division for hospitality. This work was supported by the Swedish Council for Planning and Coordination of Research.

Received 5 October 1981; accepted 3 February 1982.

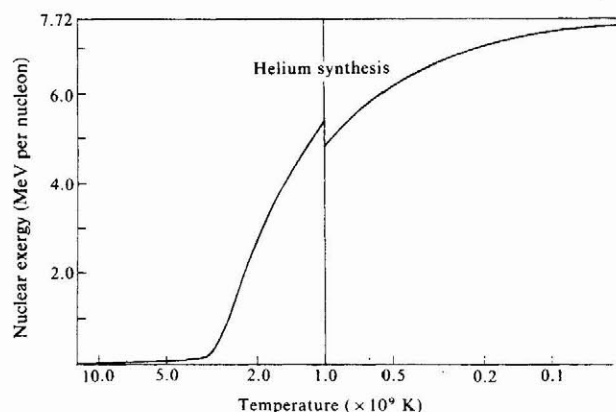


Fig. 2 Nuclear exergy as a function of the cosmic temperature T . Helium synthesis in this model takes place instantaneously at $T = 1.0 \times 10^9$ K. It involves an exergy consumption of around 0.6 MeV per nucleon, one-third of which is compensated for by the subsequent exergy production. This can be compared with the exergy loss of around 10 eV per nucleon some 700,000 yr later, when the electrons condense into atoms.

- Rant, Z. *Forsch. Geb. IngWes.* **22**, 36 (1956).
- Weinberg, S. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (Wiley, New York, 1972).
- Tolman, R. C. *Relativity, Thermodynamics and Cosmology* (Oxford University Press, 1934).
- Layzer, D. *Scient. Am.* **233**, 56 (1975).
- Landsberg, P. *Thermodynamics and Statistical Mechanics*, 233, 366-367 (Oxford University Press, 1978).
- Baehr, H. D. *Energie und Exergie* (VDI, Düsseldorf, 1965).
- Tribus, M., Shannon, P. T. & Evans, R. B. *Am. Inst. Chem. Engng J.* **12**, 244 (1966).
- Evans, R. B. thesis, Dartmouth College, Hanover, New Hampshire (1969).
- Tribus, M. & McIrvine, E. C. *Scient. Am.* **225**, 179 (1971).
- Evans, R. B. in *The Maximum Entropy Formalism*; 15 (eds Levine, R. D. & Tribus, M.) (MIT Press, 1979).
- Kullback, S. *Information Theory and Statistics* (Wiley, New York, 1959).
- Fry, J. N., Olive, K. A. & Turner, M. S. *Phys. Rev. D* **22**, 2953 (1980); **D22**, 2977 (1980).
- Kolb, E. W. & Wolfram, S. *Phys. Lett.* **91B**, 217 (1980); *Nucl. Phys.* **B172**, 224 (1980).
- Olive, K. A., Schramm, D. N., Steigman, G., Turner, M. S. & Yang, J. *Astrophys. J.* **246**, 557 (1981).
- Peebles, P. J. *Astrophys. J.* **146**, 542 (1966).
- Wagoner, R. V., Fowler, W. A. & Hoyle, F. *Astrophys. J.* **148**, 3 (1967).
- Schramm, D. N. & Wagoner, R. V. *A. Rev. nucl. Sci.* **27**, 37 (1977).
- Dyson, F. J. *Scient. Am.* **225**, 50 (1971).
- Landau, L. D. & Lifschitz, E. M. *Statistical Physics*, 340-343 (Pergamon, Oxford, 1959).
- Hoyle, F. & Fowler, W. *Nature* **197**, 533 (1963).
- Langacker, P. *Phys. Rep.* **72**, 187 (1981).
- Sakharov, A. *JETP Lett.* **5**, 24 (1967).
- Yoshimura, M. *Phys. Rev. Lett.* **41**, 281 (1978).
- Ignatiev, A., Kranikov, N., Kuzmin, V. & Tankhelidze, A. *Phys. Lett.* **76B**, 436 (1978).
- Dimopoulos, S. & Susskind, L. *Phys. Rev. D* **18**, 4500 (1979).
- Ellis, J., Gaillard, M. K. & Nanopoulos, D. V. *Phys. Lett.* **80B**, 360 (1978); **82B**, 464 (1979).
- Weinberg, S. *Phys. Rev. Lett.* **42**, 850 (1979).
- Toussaint, D., Treiman, S. B., Wilczek, F. & Zee, A. *Phys. Rev. D* **19**, 1036 (1979).
- Nanopoulos, D. V. & Weinberg, S. *Phys. Rev. D* **20**, 2484 (1979).
- Kolb, E. W. & Wolfram, S. *Phys. Lett.* **91B**, 217 (1980).
- Ellis, J., Gaillard, M. K., Nanopoulos, N. V. & Rudaz, S. *Nucl. Phys.* **B176**, 61 (1980).

Cygnus X-3 observed at photon energies above 500 GeV

R. C. Lamb*, C. P. Godfrey*, W. A. Wheaton† & T. Tümer‡

* Iowa State University, Ames, Iowa 50011, USA

† Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA

‡ University of California, Riverside, California 92521, USA

We have used the twin 11-m diameter mirrors of the NASA Jet Propulsion Laboratory's solar energy facility to observe high-energy γ rays from Cyg X-3 using the atmospheric Cerenkov technique. Observations of photons with energies above 10^{12} eV modulated with the 4.8 h period of the X-ray source have been reported previously by the Crimean Astrophysical Observatory¹⁻⁶ and the Fred Whipple Observatory^{7,8}. We report here data from $\sim 10^5$ air shower events obtained 29 August–6 September 1981 with an approximate threshold energy of 500 GeV. A positive signal with an amplitude of $10.9 \pm 2.5\%$ of the background cosmic ray rate appears near phase 0.6 of the 4.8-h cycle, in which phase 0.0 corresponds to minimum X-ray emission. This observation, together with previous Cerenkov detections, indicates that the high energy emission from Cyg X-3 is evolving on a time scale of a few years which suggests a recent origin for the system in its present form.

The JPL solar energy mirrors are located at Edwards Air Force Base, California at a dry desert location at an elevation of 0.7 km. Each mirror was viewed by a single, fast photomultiplier which gave full fields of view of 1.0° and 1.8° . (Requirements of compatibility with the solar research dictated the unequal size tubes.) The mirrors are separated by 24 m. The rate of air shower coincidences between the two mirrors is $\sim 60\%$ of the rate for a single mirror. Preliminary tests with the mirrors in June–July 1981 gave a background coincidence rate of $\sim 4 \text{ s}^{-1}$ for air shower events near the zenith. However, before our August–September run, the focus of one of the mirrors was changed for solar energy purposes to reduce the effective size of the mirror to $\sim 5 \text{ m}$. The data reported here were obtained in conditions in which the background rate at the zenith was about 1 s^{-1} .

Extensive alignment tests and angular calibrations were performed with each mirror using the image information from paraxially mounted television cameras. During observations the pointing of each mirror was monitored and found to be reproducible to better than 0.1° . Data associated with air shower events were recorded magnetically when signals from the two detectors occurred within 50 ns of each other. The pulse height associated with each detector, the time difference between the two signals, and the universal time to 0.1 ms were recorded. In addition singles and accidentals rates were monitored. The accidental rate for most scans was $< 1\%$. Analysis of the data offline showed that coincident events were characterized by a peak whose FWHM was $\sim 10 \text{ ns}$, equivalent to a resolving time of $\sim 5 \text{ ns}$. For the data reported here we have selected events with a 'tight' coincidence requirement corresponding to a resolving time of 3.8 ns.

Observations were made in drift scans which typically lasted 33 min, starting 17 min before and ending 16 min after transit. Following the procedure of Danaher *et al.*⁷, we divided the data from an individual drift scan into three regions: two background regions more than 5 min before and after source transit, and an on-source region. We used an on-source region of 7 min centred at the source transit time. From a sum of the drift scans we determined that the 2nd mag star, γ Cygni, which transited 10 min before Cyg X-3, increased the background rate by $3 \pm$

1%; therefore we have deleted data during an interval of 6 min about γ Cyg from our analysis.

The data rate was analysed in 30-s bins. A scan was rejected if: (1) any 30 s interval gave a $\pm 4\sigma$ departure from the mean; (2) the two background observations failed to agree at the 5% level; (3) a χ^2 uniformity test for the run as a whole failed at the 5% level. In addition a Poisson homogeneity test at the 5% level was applied to the data binned in 10 s bins. Of 95 drift scans recorded during the 9 day run, 81 passed all tests.

Figure 1 shows the ratio of the on-source counting rate to the background counting rate corresponding to various phases of the 4.8-h periodicity. Errors have been calculated both experimentally and on the basis of Poisson statistics. For the scans which passed the selection tests there is no significant difference, and the Poisson values are used. For our phase calculation we have used the X-ray parameters of van der Klis and Bonnet-Bidaud⁹. Their quadratic and linear parameters give a difference in phase of 0.04 at our epoch. We have used the quadratic parameters tied to their well determined T_{min} of May–June 1980. If the Crimean values⁶ are used then the phases of Fig. 1 are increased by 0.115.

Eight of the 10 phase bins are consistent with no signal. However, in the adjacent phase bins from 0.5 to 0.7 there are 3.1 and 2.9 σ excesses, which together constitute a 4.2 σ excess which we interpret as a positive detection. If we bin the data in the signal region in bins of 0.05 phase, the ratios and significances for the four bins beginning at 0.5 phase are: 1.04 (0.7 σ), 1.12 (3.2 σ), 1.09 (2.2 σ), and 1.08 (1.9 σ). The 23 scans corresponding to phases between 0.5 and 0.7 have been added together and their sum is shown in Fig. 2a. These data have been fitted to a level background plus a gaussian of variable width and height centred at the transit time. The best fit value of the height is $10.9 \pm 2.5\%$ of the background, equivalent to a 4.4 σ detection. Figure 2b shows the sum of the scans with phases outside of the range 0.5–0.7. The horizontal line corresponds to the average background rate. The flux averaged over a cycle corresponding to the effect is $\sim 8 \times 10^{-11} \text{ photons cm}^{-2} \text{ s}^{-1}$ in reasonable agreement with previous detections at higher energies.

The phase of the observed emission, 0.6 ± 0.1 , agrees with the phase typically seen for maximum X-ray emission⁹ and agrees approximately with the 1980 observation of Danaher *et al.*⁷, after accounting for the different methods of phase calculation.

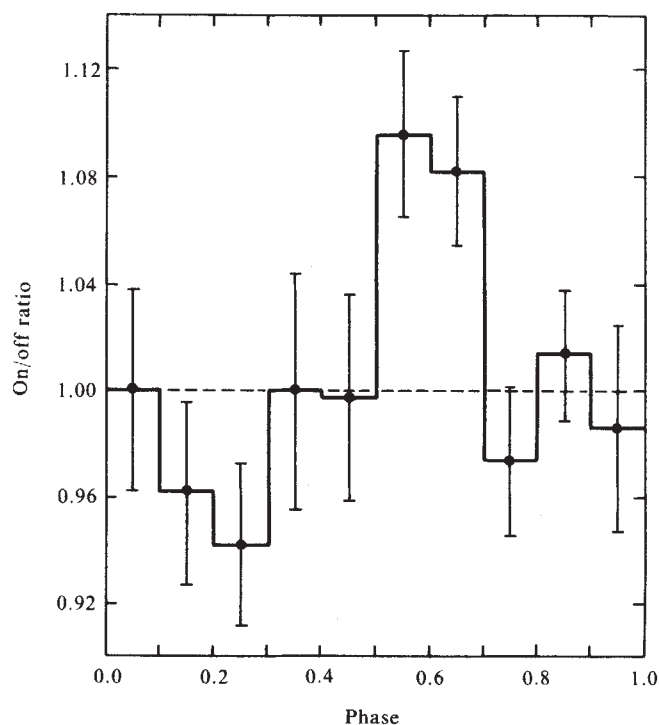


Fig. 1 Phase histogram of the γ -ray emission from Cyg X-3.

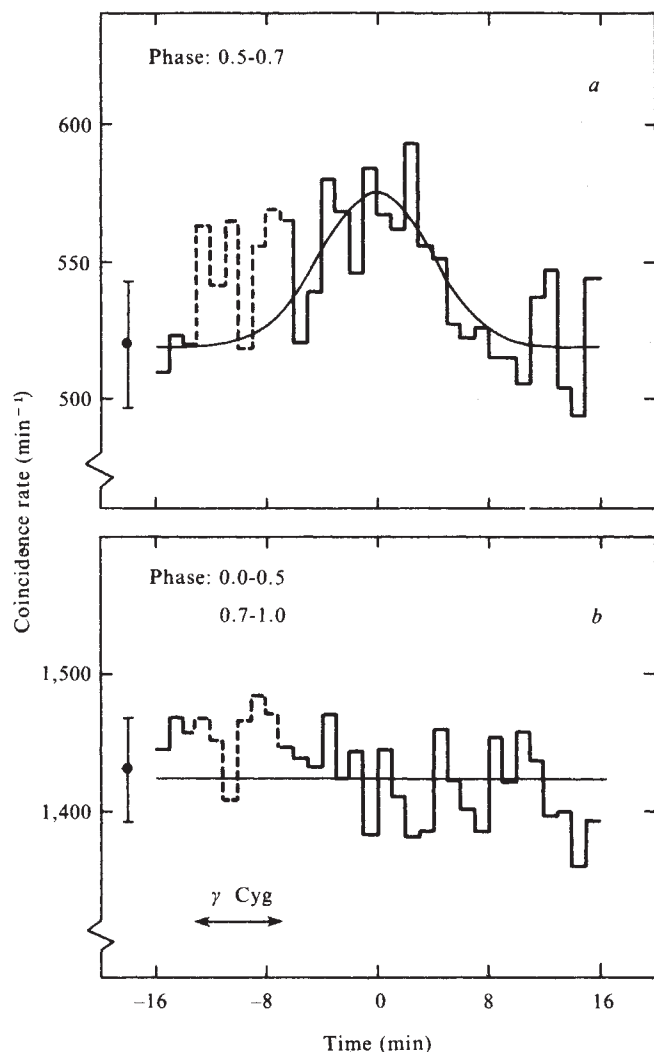


Fig. 2 Coincidence rate in c.p.m. versus drift scan time. The transit of Cyg X-3 occurs at time 0. *a*, The sum of the 23 scans corresponding to phases between 0.5 and 0.7. The smooth curve is the best fit of a gaussian, centred at the transit, plus a level background. The height of the gaussian is $10.9 \pm 2.5\%$ of the background, equivalent to a 4.4σ detection. *b*, The sum of the remaining scans. The horizontal line corresponds to the average background rate. For these phases there is no evidence for γ -ray emission from Cyg X-3. The dotted portion of each histogram corresponds to the 6-min interval in which the 2nd mag star, γ Cyg, is in the field of view. These data were not used in any of the analyses.

tion⁸. However, the width of ~ 0.2 cycles seen here may be inconsistent with their results and is inconsistent with the Crimean results which have been interpreted as indicating activity confined to 0.05 cycles^{4,6}. Also, we see no evidence for emission at other phases as observed in some of the earlier Crimean measurements. The disparity in these observations may indicate changes in the characteristics of the high energy emission on a time scale of a few years. These apparent changes and the apparent 100 MeV variability implied by the SAS 2 detection¹⁰ and the COS B upper limit¹¹ support the view that Cyg X-3 is a system which is rapidly evolving, and that it may have been created in its present form in the recent past. As has been pointed out^{12,13}, the presence of high energy γ rays favour models with a young pulsar in rapid rotation. If the pulsed fraction increases with energy as in the case of the Crab¹⁴ and Vela¹⁵ pulsars, then future high statistics ground based measurements may be the most promising method of detection.

We thank the staff of JPL Edwards Test Site for cooperation and especially John Woodbury, Terry Hagen, and Jack Whitehead. W.A.W. was supported by an NRC Resident Associate-

ship during part of this work. This work was supported partially by the US Department of Energy, NASA, and a grant from the California Institute of Technology President's Fund.

Received 11 January; accepted 19 February 1982.

1. Vladimirov, B. M., Stepanian, A. A. & Fomin, V. P. *Proc. 13th int. Conf. Cosmic Rays* **1**, 456-460 (1973).
2. Vladimirov, B. M., Neshpor, Yu. I., Stepanian, A. A. & Fomin, V. P. *Proc. 14th int. Conf. Cosmic Rays* **1**, 118-122 (1975).
3. Stepanian, A. A., Vladimirov, B. M., Neshpor, Yu. I. & Fomin, V. P. *Proc. 15th int. Conf. Cosmic Rays* **1**, 135-140 (1977).
4. Neshpor, Yu. I. *et al. Astrophys. Space Sci.* **61**, 349-355 (1979).
5. Mukanov, J. B., Nesterova, N. M., Stepanian, A. A. & Fomin, V. P. *Proc. 16th int. Conf. Cosmic Rays* **1**, 143-146 (1979).
6. Neshpor, Yu. I., Zyskin, Yu. L., Mukanov, J. B., Stepanian, A. A. & Fomin, V. P. *Phil. Trans. R. Soc. A* **301**, 633-634 (1981).
7. Danaher, S., Fegan, D. J., Porter, N. A. & Weekes, T. C. *Nature* **289**, 568-569 (1981).
8. Weekes, T. C., Danaher, S., Fegan, D. J. & Porter, N. A. *Astr. Astrophys. Lett.* **104**, L4-L6 (1981).
9. van der Klis, M. & Bonnet-Bidaud, J. M. *Astr. Astrophys. Lett.* **95**, L5-L7 (1981).
10. Lamb, R. C., Fichtel, C. E., Hartman, R. C., Kniffen, D. A. & Thompson, D. J. *Astrophys. J. Lett.* **212**, L63-L66 (1977).
11. Bennett, K. *et al. Astr. Astrophys.* **59**, 273-274 (1977).
12. Bignami, C. F., Maraschi, L. & Treves, A. *Astr. Astrophys.* **55**, 155-158 (1977).
13. Milgrom, M. & Pines, D. *Astrophys. J.* **220**, 272-278 (1978).
14. Thompson, D. J., Fichtel, C. E., Hartman, R. C., Kniffen, D. A. & Lamb, R. C. *Astrophys. J.* **213**, 252-262 (1977).
15. Kanbach, G. *et al. Astr. Astrophys.* **90**, 163-169 (1980).

Characterization of radioactive fallout from pre- and post-moratorium tests to polar ice caps

Minoru Koide, Robert Michel & Edward D. Goldberg

Scripps Institution of Oceanography, La Jolla, California 92093, USA

Michael M. Herron & Chester C. Langway Jr

State University of New York at Buffalo, Amherst, New York 14214, USA

Polar glaciers contain a detailed historical record of artificial radioactive fallout from the atmosphere¹⁻⁴. The records are continuous and time frames can be introduced with uncertainties of about ± 1 yr. We have previously indicated that there may be characteristic nuclide compositions of fallout produced by weapons testing by different countries from records in ice sheets. For example, the testing period dominated by the US activity in the early 1950s produced a markedly low $^{238}\text{Pu}/^{239+240}\text{Pu}$ ratio and a very high $^{241}\text{Pu}/^{239+240}\text{Pu}$ ratio compared with those resulting from tests dominated by the USSR activity in the early 1960s^{1,3,5}. Here we use analyses on a recently collected Greenland ice sheet core to extend these results and find the characteristic $^{239+240}\text{Pu}/^{137}\text{Cs}$ and $^{239+240}\text{Pu}/^{90}\text{Sr}$ ratios for these two periods. The atmospheric behaviours of tritium compared with those of other artificial radionuclides are recorded in its fluxes to the polar ice sheets.

From 40-m deep, 7.5-cm diameter firn cores were recovered from Dye-3 ($65^{\circ}11' \text{N}$, $43^{\circ}50' \text{W}$) in June 1980 as part of the joint US-Danish-Swiss Greenland Ice Sheet Program. Density measurements and visible stratigraphical observations were made on each core section. The cores were then stratigraphically correlated and combined to form individual samples for transuranic and fission product analyses representing 2 yr on average. ^3H analyses were performed on the yearly samples. The analytical techniques have been described elsewhere¹.

The average accumulation rate, based on ^{210}Pb profile in the glacier, is $52 \text{ g H}_2\text{O cm}^{-2} \text{ yr}^{-1}$ (Fig. 1a). The inventory of ^{210}Pb , corrected for decay to the time of collection, to the Dye-3 site (9.8 mCi km^{-2} for the period 1944-80) is similar to that of South Dome, Greenland ($63^{\circ}31.6' \text{N}$, $44^{\circ}34.5' \text{W}$), for the period 1945-75 that we measured previously (10.2 mCi km^{-2} , ref. 1). A lower ^{210}Pb inventory of 4.4 mCi km^{-2} is computed

for the Crete, Greenland, Site at 71°07' N and 37°19' W (ref. 6).

The ^{210}Pb derived accumulation rate compares well with the long-term average rate of $49 \text{ g cm}^{-2} \text{ yr}^{-1}$ determined from stable oxygen isotope analysis⁷. The samples cover on average two year intervals with a range of 1–3 yr, based on the ^{210}Pb ages (Fig. 1).

The plutonium inventories to the northern and southern hemispheric polar sites indicate a higher contribution from the US dominated pre-moratorium atmospheric nuclear tests compared with the USSR dominated post-moratorium atmospheric nuclear tests (Fig. 1; Table 1) than previously estimated. The moratorium period was from November 1958 to September 1961. On the basis of the total atmospheric bomb yields of these two periods, the ratio⁸ of post- to pre-moratorium testing is around 70:30. On the basis of the accumulated amounts of

plutonium in both Arctic and Antarctic ice sheets, the contributions from the pre-moratorium tests varies between 41 and 64%. The value from three of the four sites clusters about 43%.

On the other hand, the ^{137}Cs and ^{90}Sr results (Table 1) accord with the 70:30 ratio from the bomb yields. This was expected because ^{137}Cs and ^{90}Sr are fission products whereas Pu is not. Pu production does not necessarily correlate with total bomb production as it is also dependent on the type of weapon construction.

The integrated fluxes of ^{238}Pu and $^{239+240}\text{Pu}$ are about a factor of two lower in the Dye-3 Greenland sites compared with those of the South Dome Greenland site. The fluxes are: ^{238}Pu , 0.003 and 0.006 mCi km^{-2} ; and $^{239+240}\text{Pu}$, 0.27 and 0.7 mCi km^{-2} . This situation may arise from unique meteorologies associated with the delivery of stratospheric fallout to the two areas, only

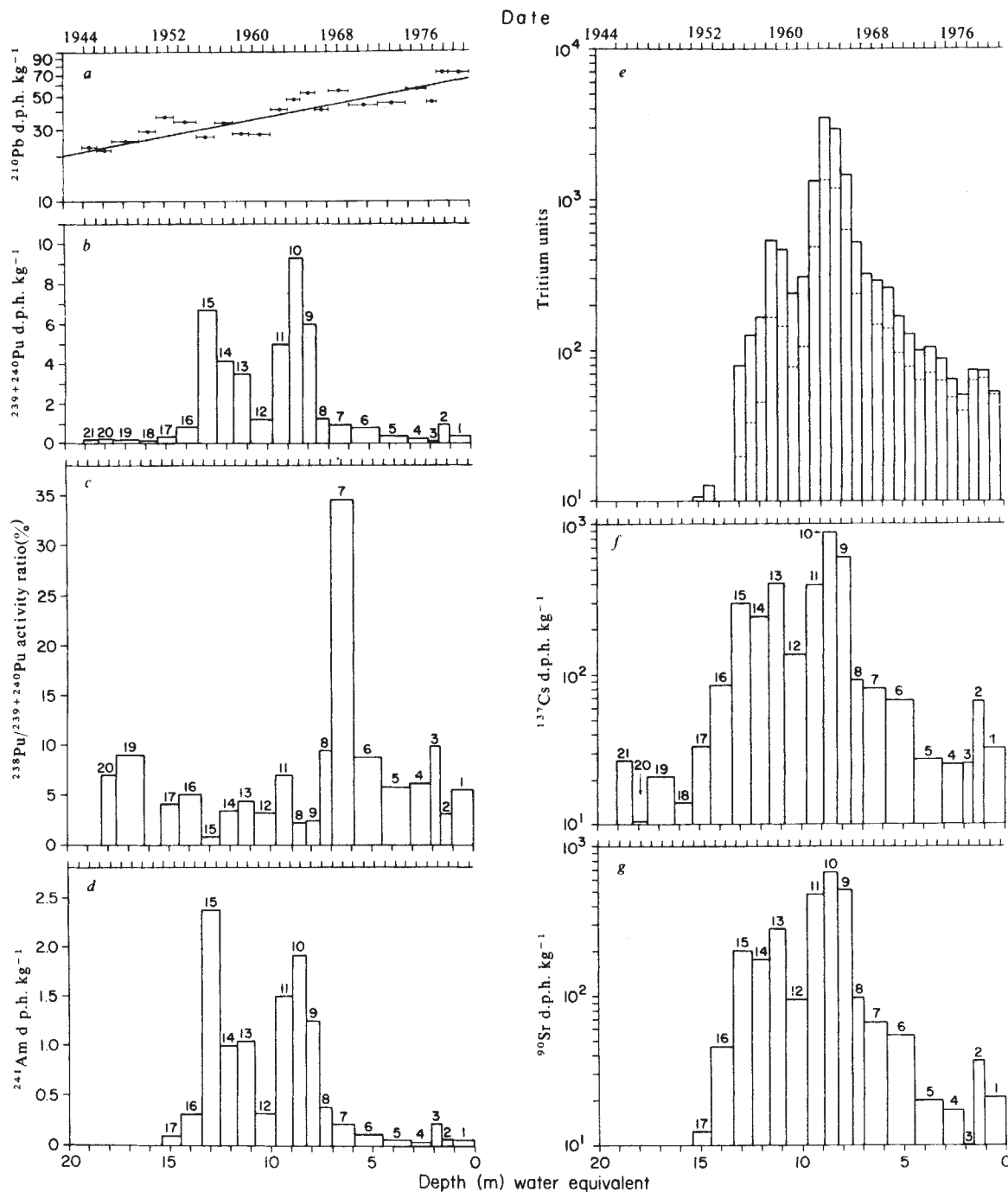


Fig. 1 Radionuclides in the Greenland Ice Sheet at Dye-3 Station. The time scale is derived from the ^{210}Pb profile (a). The ^{238}Pu (c), ^3H (e), ^{137}Cs (f) and ^{90}Sr (g) concentrations are corrected to the time of deposition. The tritium activity observed is indicated by the dashed line in the histogram (e). The average counting errors and their ranges (in parentheses) are: $^{239+240}\text{Pu}$, 6.8% (2.9–15); ^{238}Pu , 24% (8–50), intervals from SNAP-9A input 8% (8–14); ^{241}Am , 10.4% (4.7–19). For ^{210}Pb , ^{137}Cs and ^{90}Sr , the counting errors were <3%.

a few hundred kilometres apart. This result suggests the site-specificity of the fluxes and the uncertainty in utilizing them as representative of a large area. Lambert *et al.*⁹ found a similar situation with the distribution of ^{90}Sr in the Antarctic ice-sheet.

The lowest $^{238}\text{Pu}/^{239+240}\text{Pu}$ activity ratio occurs in Level 15, 0.7% (Fig. 1c). This is concordant with the low values of the ratio (<1%) found at the Antarctic sites, J-9 and Dome C, for similar times^{1,3}.

The Arctic ratios of the integrated fluxes ($^{239+240}\text{Pu}/^{137}\text{Cs}$, $^{239+240}\text{Pu}/^{90}\text{Sr}$ and $^{137}\text{Cs}/^{90}\text{Sr}$; Table 1) are remarkably similar for the pre- and post-moratorium intervals to those of the Antarctic site at J-9. These ratios distinctly characterize the pre- and post-moratorium testing: $^{239+240}\text{Pu}/^{137}\text{Cs}$, pre-moratorium 0.016 and 0.019, post-moratorium 0.012 and 0.012; $^{239+240}\text{Pu}/^{90}\text{Sr}$, pre-moratorium 0.023 and 0.027, post-moratorium 0.013 and 0.011; $^{137}\text{Cs}/^{90}\text{Sr}$, pre-moratorium 1.6 and 1.5, post-moratorium 1.1 and 0.9. Note that an average post-moratorium activity ratio of $^{239+240}\text{Pu}/^{137}\text{Cs}$ of 0.012 was obtained from the annual atmospheric data of Richland, Washington (1962–75), New York (1969–75) and Harwell, UK (1972–75)¹⁰. There are no data available for comparison with our pre-moratorium ratios.

The high Pu/Cs and high Pu/Sr ratios for the US dominated pre-moratorium testing period in the 1950s in the Antarctic samples¹ are confirmed here for the Dye-3 samples. Levels 14 and 15 (Fig. 1) correspond to this testing period whose fallout in the Arctic took place primarily between 1955 and 1958. The ratios have values during this period twice those of previous and subsequent years. On the other hand, the Cs/Sr ratio during this period is significantly higher than during the post-moratorium period (Table 1). The weapons debris, introduced to the stratosphere, apparently becomes mixed there in times that are short with respect to a residence time for the particles of the order of 1 yr. Clearly, the polar glaciers will record the fallout without strong distortions due to meteorology or differences in atmospheric chemistry among the different elements. Thus, as well as the low values of the $^{238}\text{Pu}/^{239+240}\text{Pu}$ ratio in the US dominated 1950s testing period we have identified another characteristic of the radioactive debris at this time.

Again, as in the Antarctic ice sheet, the recorded times of arrival of the plutonium nuclides, the caesium and strontium isotopes and the tritium follow a pattern where the transuranics are the first to accumulate, followed by the alkali and the alkaline earth nuclides and then the tritium. The percentages of the contents of nuclides in stratum 15 relative to the total in strata 13, 14, 15 and 16 are: $^{239+240}\text{Pu}$, 46; ^{137}Cs or ^{90}Sr , 31; and ^3H , 17 (Fig. 1). Our original explanation¹ indicated that the particulate transuranic nuclides have the shortest residence times in the stratosphere and the tritium, in the form of water, has the longest. ^{137}Cs and ^{90}Sr have intermediate values of the stratospheric residence times. This may relate to their hydro-

philic natures as ions and the stratospheric times involved in their abilities to form an association with waters.

The tritium concentrations in precipitation (Fig. 1e) show a pattern similar to that of ^{90}Sr and ^{137}Cs . Two distinct peaks are evident corresponding to fusion weapons testing in the mid-1950s and 1961–62. The deposition of tritium in the pre- and post-moratorium time periods is also similar to that of ^{90}Sr and ^{137}Cs . Roughly 85% of the total tritium deposition occurred in the post-moratorium period at both Dye-3 and J-9. This is due to the high production of tritium from the detonation of fusion devices¹¹.

Figure 2a displays the deposition at Dye-3 relative to that at J-9 on a year by year basis. Four major events are evident from this graph. All events indicate a sudden increase of deposition at Dye-3 relative to J-9. Note that the tritium deposition usually increases at both stations at about the same time, but the increase for Dye-3 is larger. The first peak follows the beginning of fusion bomb testing in the mid-1950s. The second event follows the large test series of 1961–62. The two smaller events seem to correlate with nuclear series of France and the People's Republic of China following the 1963 moratorium. These tests indicate that for all major series, the initial tritium deposition increased most in the northern polar region. After 1–2 yr, the extreme differences in deposition appear to approach a baseline ratio (Dye-3/J-9 deposition) of slightly greater than six. This is close to the difference in the annual precipitation at the two sites (52 cm compared with 8.9 cm) indicating very little difference in the specific activity at the two sites within about 2 yr after a test series. The small difference in specific activity could be accounted for by the higher tritium concentrations found in Northern Hemisphere surface waters¹².

This pattern can be explained as follows. Most nuclear testing produced a large increase in the tritium burden of the northern hemisphere's troposphere and stratosphere. The tropospheric tritium quickly deposited on the Earth's surface while the stratospheric tritium began to mix into the water vapour of the northern troposphere and southern stratosphere. After the initial tropospheric input was deposited, the deposition at Dye-3 was primarily from stratospheric input, although some tritium from re-evaporation was also present. The southern polar troposphere received a much smaller initial tritium input from the nuclear tests. Most of the tritium it received came from the stratosphere. Thus, in less than 1 yr, tritium concentrations at both polar tropospheres were dependent on stratospheric input. From Fig. 2, it appears that transient stratospheric inputs level off in ~2 yr. Due to the short residence time of tritium in the atmosphere, deposition at Dye-3 was ~20 times larger than that at J-9.

Figure 2b, c shows the relative deposition of tritium versus ^{90}Sr and ^{137}Cs at both sites. As expected, very little tritium is present relative to these two isotopes in the early nuclear tests.

Table 1 Inventories of artificial radionuclides in polar ice sheets

	Ref.	$^{239+240}\text{Pu}$	^{137}Cs	^{90}Sr	$^3\text{H}^*$	$^{239+240}\text{Pu}$	$^{239+240}\text{Pu}$	^{137}Cs
		Inventory (%)	Inventory (%)	Inventory (%)	Inventory (%)	^{137}Cs	^{90}Sr	^{90}Sr
Arctic								
Dye-3 Greenland (65°11' N, 43°50' W)	This work	0.3	21.3	16.7	2.27×10^{-3}			
Post-moratorium		59	65	72	87	0.012	0.013	1.1
Pre-moratorium		41	35	28	13	0.016	0.023	1.6
South Dome, Greenland (63°31.6' N, 44°34.5' W)	2	0.7						
Post-moratorium		56						
Pre-moratorium		44						
Antarctic								
J-9 (82°22' S, 168°40' W)	1	0.04	2.8	2.7	1.12×10^{-4}			
Post-moratorium		57	68	77	85	0.012	0.011	0.9
Pre-moratorium		43	32	23	15	0.019	0.027	1.5
Dome C (74°39' S, 123°10' E)	3	0.04						
Post-moratorium		36						
Pre-moratorium		64						

The inventories calculated for all nuclides except $^{239+240}\text{Pu}$ were corrected for decay to the time of deposition as derived from ^{210}Pb chronology and are in mCi km^{-2} .
* ^3H inventory units in g km^{-2} .

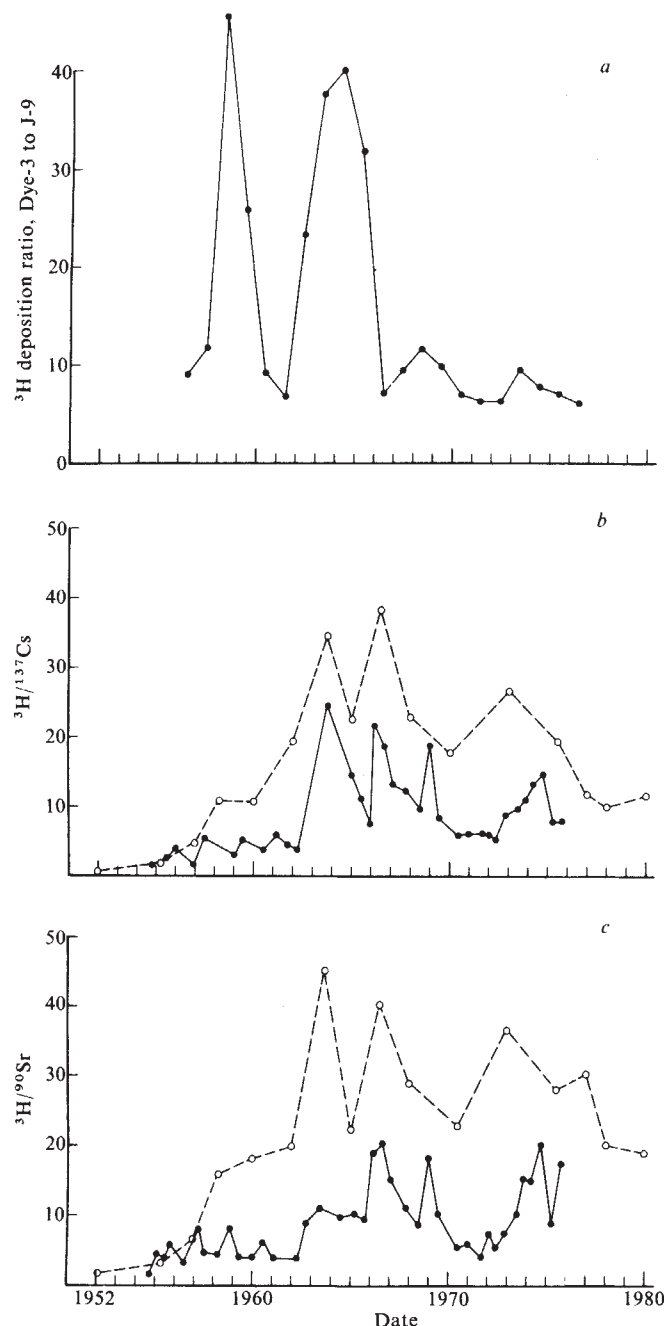


Fig. 2 a, Ratio of ^3H deposition at Dye-3 to J-9 as a function of time; b, $^3\text{H}/^{137}\text{Cs}$ ratio as a function of time at the Greenland and Antarctic site; c, $^3\text{H}/^{90}\text{Sr}$ ratio as a function of time at the Greenland and Antarctic site. $\circ$, Dye-3; $\bullet$, J-9.

Following the advent of fusion testing, tritium deposition increased.

These results emphasize the need to carry out similar measurements on mid-latitudinal (alpine) glaciers. We suspect that such glaciers, which exist at all latitudes of the Earth, may hold a detailed record of global artificial radioactivity fallout not found elsewhere and which should contain surprises. With the development of methods over the past decade or so to measure such nuclides at ^{239}Pu , ^{240}Pu , and ^{241}Pu , ^{242}Cm , and the fission nuclides, their time and space distribution in glaciers should be most revealing about the nature of the weapons and the meteorology of transport. We recognize the difficulties in the identification of appropriate sampling sites where summer melting is negligible or non-existent such that percolation does not smear the record. Approximately 8% of each year's snowfall at Dye-3 was melted and refrozen as ice layers¹³. This suggests

that moderate melting at mid-latitude glacier sites would not seriously perturb the record of radionuclide fallout. We submit that the search for such a record will be rewarding.

This work was supported by grants from the Department of Energy, Environmental Programs, Division of Biomedical and Environmental Research and the National Science Foundation, Division of Polar Programs. H. Ruffli drilled the firn cores. We thank Dr T. M. Beasley of Oregon State University for critical comments.

Received 30 November 1981; accepted 13 February 1982.

1. Koide, M., Michel, R., Goldberg, E. D., Herron, M. M. & Langway, C. C. *Earth planet. Sci. Lett.* **44**, 205 (1979).
2. Koide, M., Goldberg, E. D., Herron, M. M. & Langway, C. C. *Nature* **269**, 137 (1977).
3. Cutter, G. A., Bruland, K. W. & Risebrough, R. W. *Nature* **279**, 628 (1979).
4. Jouzel, J., Pourchet, M., Lorius, C. & Merlivat, M. *Earth planet. Sci. Lett.* **45**, 188 (1979).
5. Koide, M. & Goldberg, E. D. *Earth planet. Sci. Lett.* **54**, 239 (1981).
6. Goldberg, E. D. in *Radioactive Dating* 121-131 (IAEA, Vienna, 1963).
7. Reeh, N. et al. *J. Glaciol.* **20**, 27 (1978).
8. Carter, W. & Moghissi, A. A. *High Phys.* **33**, 55 (1977).
9. Lambert, G., Ardavin, B. & Brichet, E. *Earth planet. Sci. Lett.* **11**, 317 (1971).
10. Perkins, R. W. & Thomas, C. W. in *Transuranic Elements in the Environment* (ed. Hanson, W. C.) 53 (Technical Information Center, US Department of Energy, 1980).
11. Erickson, E. *Tellus* **15**, 303 (1965).
12. Burton, J. D. in *Chemical Oceanography* Vol. 3, Chap. 18 (Academic, New York, 1975).
13. Herron, M., Herron, S. L. & Langway, C. C. *Nature* **293**, 389 (1981).

Fission product transmutation effects on high-level radioactive waste forms

W. J. Gray

Materials Department, Pacific Northwest Laboratory, Richland, Washington 99352, USA

The properties of potential nuclear waste forms have been under investigation for more than a decade to ensure their safe handling, packaging, transport, storage, and disposal. Two kinds of high-level radioactive waste (HLW) have received most attention: (1) that which could be produced from reprocessed commercial reactor fuel, and (2) the slightly less radioactive waste from military activities. Both kinds would be converted to solid, nondispersible forms such as glass or crystalline ceramics for disposal. After disposal, self-radiation damage, primarily radiation-induced atomic displacements, can change the properties of these waste forms and possibly increase the chances that radioactivity could be released to the environment. β decay of the fission products causes some displacements but α decay, and the associated recoil nuclei, of the actinide elements produces several orders of magnitude more displacements. Most recent studies on radiation effects have therefore concentrated on α decay¹⁻³. A potentially more important aspect of β decay is the associated transmutation of fission products which usually results in changes of both valence and ionic size, both of which may not be readily accommodated by HLW solids. ^{137}Cs and ^{90}Sr are the two main sources of transmutations, but they have not been studied because of their ~ 30 -yr half lives. The problem has now been circumvented by preparing simulated waste forms and other solids containing ^{134}Cs , which has only a 2-yr half life. This study will not be completed for at least another year, but we report here that results obtained after 2 yr of decay are encouraging in that negligible changes have been observed in the properties of all the materials studied.

Fission products would comprise 40-90% of the HLW stream; the rest would consist of actinides and reprocessing elements. Most ($\sim 84\%$) of the fission products are totally non-radioactive and about half of the radioactive isotopes have half lives $\geq 10^5$ yr. Only $\sim 2\%$ have half lives ≤ 1 yr. Thus, ^{137}Cs and ^{90}Sr , which comprise $\sim 6\%$ of the fission products, are the two isotopes of major concern with regard to transmutations. Caesium decays to barium which results in a change in valence from +1 to +2 and a reduction in ionic radius of 20%. Strontium

Table 1 Density of HFIR samples and non-radioactive standards

No.	Material Type	Wt % Cs	Density (g cm ⁻³)*		
			June 1979	February 1980	June 1981
3S	Standard	—	3.209±0.007	3.203±0.004	3.214±0.007
5S	Standard	—	2.689±0.009	2.690±0.000	2.691±0.007
1	Simulated waste glass	2.63	1.993±0.018	1.990±0.010	1.992±0.004
2	High caesium glass with ²⁴⁴ Cm	28.8	2.704±0.015	2.727±0.007	2.731±0.005
3	High caesium glass	31.1	2.835±0.021	2.857±0.018	2.911±0.028
4	High silica glass	3.65	2.362±0.005	2.360±0.000	2.381±0.015
5	High alumina glass	3.27	2.315±0.004	2.310±0.000	2.309±0.007
7	Pollucite	42.7	2.970±0.020	2.970±0.000	2.994±0.003
8	Supercalcine	3.16	4.207±0.022	4.208±0.010	4.199±0.010

* Errors listed represent twice the standard deviation of the mean.

decays to zirconium; thus the valence changes from +2 to +4 and the reduction in ionic radius is ~29%. However, both ¹³⁷Cs and ⁹⁰Sr have half lives of ~30 yr, so it is not practical to use them to study the effects of transmutations.

Two possible ways of studying the effects of transmutations are: (1) to prepare a simulated waste using particular short-lived radioactive isotopes; and (2) to prepare a waste form with nonradioactive material and then irradiate it with neutrons to generate suitable short-lived isotopes. The first method would be prohibitively expensive; the second method is feasible with caesium but not strontium because naturally occurring isotopes of the latter have quite small neutron capture cross-sections. In fact, ¹³⁴Cs is the only suitable isotope that can be generated by neutron irradiation. All other elements have neutron capture cross-sections that are too small, the half lives of the radioisotopes produced are too short or too long, or the decay scheme does not yield suitably large changes in valence and ionic size.

Natural caesium is 100% ¹³³Cs, which has a fairly large thermal neutron capture cross-section. Therefore, a reasonable fraction of the caesium can be converted through neutron irradiation to ¹³⁴Cs, which decays to barium with a 2.06-yr half life. However, because ¹³⁴Cs also has a large thermal neutron capture cross-section that yields the very long-lived ¹³⁵Cs isotope, the amount of ¹³⁴Cs produced reaches a maximum rather than continuing to increase with radiation dosage. The high flux isotope reactor (HFIR) at Oak Ridge National Laboratory, Oak Ridge, Tennessee, was the most efficient reactor available to generate ¹³⁴Cs because of its favourable neutron

energy spectrum and very high neutron flux. Calculations showed that a maximum of 20% of the caesium would be converted to ¹³⁴Cs during an irradiation period of ~100 days.

Seven glasses and ceramics containing caesium were irradiated in the HFIR during the first 3 months of 1979. Immediately following irradiation, all samples were annealed to remove the neutron radiation damage effects that could mask changes caused by subsequent transmutations. The glass materials were remelted at 900–1,000 °C for 30 min, and the crystalline materials were annealed at 1,150 °C for 1 h. Remelting the glasses should remove atomic displacement damage caused by the high-energy neutrons. Whether the crystalline materials were adequately annealed is not known; it was hoped that X-ray diffraction data would answer that question but those results have been unsatisfactory until quite recently.

Calculations showed, and post-irradiation measurements confirmed, that most of the activity in the irradiated samples was due to ¹³⁴Cs. Thus, any changes in properties are due to Cs → Ba transmutations with only a negligible contribution from other radioisotopes.

Post-irradiation measurements also showed that only 12% of the caesium was converted to ¹³⁴Cs rather than the 20% that had been calculated; however, this should be enough ¹³⁴Cs to observe any property changes as the caesium transmutes to barium. Post-irradiation analyses were completed immediately after the samples were annealed and are being repeated at intervals of ~1 yr. X-ray diffraction, optical and scanning electron microscopy, leach tests, and density measurements are

Table 2 Leach values of materials irradiated in HFIR

Material no.	% Of element leached in 7 days at 90 °C in deionized water*					
	Boron			Silicon		
	May 1979	April 1980	August 1981	May 1979	April 1980	August 1981
1	ND†	0.93±0.17	1.61±0.22	0.57±0.36	1.00±0.12	1.36±0.31
2	23.6±17.6	15.0±6.6§	22.7±3.0	7.8±3.4	5.7±2.3§	10.3±0.4
3	27.6±2.6	14.5±2.8§	24.3±8.9	7.0±0.7	6.5±1.9§	7.6±0.6
4	ND	0.46±0.20	0.61±0.07	0.13±0.00	0.15±0.03	0.12±0.03
5	ND	0.31±0.33	0.31±0.15	0.43±0.22	0.61±0.19	0.57±0.22
7	—‡	—	—	0.70±0.30	1.15±0.06	1.25±0.12
8	—	—	—	3.1±0.3	2.33±0.07§	2.97±0.14
	Molybdenum			Caesium		
	May 1979	April 1980	August 1981	May 1979	April 1980	August 1981
	0.11±0.04	ND	0.62±0.26	0.52±0.12	0.89±0.19	1.56±0.22
	16.5±15.8	9.5±3.6§	14.1±3.1	15.3±9.5	5.99±2.21§	16.5±3.6
	18.6±0.9	8.8±1.9§	16.4±7.5	29.9±1.3	13.8±2.8§	28.6±6.2
	—	—	—	0.36±0.03	0.38±0.02	0.58±0.07
	—	—	—	0.53±0.24	0.34±0.10	0.38±0.14
	—	—	—	1.25±0.51	1.68±0.36	2.58±0.34
	0.71±0.11§	0.09±0.02	0.08±0.04	2.95±0.16	0.57±0.26§	1.98±0.65

* Errors listed represent twice the standard deviation of the mean.

† ND, —, element not detected.

‡, Element not present.

§ Results that are out-of-line with related results and are probably in error.

being made to determine changes as the Cs $\rightarrow$ Ba transmutations occur.

Table 1 lists the materials irradiated, the amount of caesium in each and results of the density measurements. Density results for two non-radioactive reference materials are also included. Three to five measurements were made on each sample. Densities have increased slightly for four of the materials including pollucite. The largest increases are for the two high-caesium glasses, particularly the one without curium. Increases for the other two materials (4 and 7) are statistically significant, but just barely; additional measurements a year or so from now will be required definitely to establish a trend. No measureable change has been found for materials 1, 5, and 8.

Leach results are listed in Table 2. Three samples of each material were leach-tested in May 1979, three different samples of each were leached during April 1980 and so on. The leach tests were done at 90 °C in deionized water for 7 days. Water volume to sample surface area ratios were 30–40 cm; more importantly, this ratio was maintained at a constant value in each leach test of a given material. Following the leach tests, the leachate solutions were analysed for ^{134}Cs activity using a high-resolution γ -ray spectrometer which was calibrated with a ^{134}Cs standard traceable to the National Bureau of Standards.

After the leachate solutions were analysed for caesium activity, the level of radioactivity was reduced by passing the solutions through ion exchange columns to remove the caesium. The solutions were then analysed by inductively coupled plasma spectroscopy for most other cations except alkali and alkaline earths whose concentrations were altered by the ion exchange resins.

Except for the alkali and alkaline earth interferences noted above, only those elements listed in Table 2 were found in the leachate solutions in concentrations above the detection limits. Several results from the second test and one from the first test seem to be in error. If so, then the leach rates of samples 2, 3, 4, and 5 have not changed noticeably. Sample 8 also has not changed much except for a very small decrease in the leach rate of Cs. The leach rates of samples 1 and 7 have increased slightly. However, the increases are relatively unimportant from a waste management point of view as the changes are less than a factor of 10.

Interestingly, density changes were largest for samples 2 and 3 whereas their leach rates have not changed. This is probably because their leach rates are so large to start with that small changes in microstructure makes little difference.

Optical micrographs were taken of the same areas of all seven samples during each of the three post-irradiation examinations. Scanning electron micrographs were also taken during the second round of examinations and compared with electron microprobe results obtained during the first post-irradiation examination. No change in microstructure was found in any of the samples.

X-ray diffraction results have been unsatisfactory until recently when a new shielded diffractometer became operational at PNL. Only the irradiated pollucite has been examined so far, and even those results are quite preliminary. The pollucite still seems to be at least partially crystalline and the lattice parameters have not changed appreciably. These cursory results obtained about 2 yr after irradiation are thus in approximate agreement with electron diffraction results obtained on a pollucite sample about 6 months after irradiation. The new X-ray equipment will allow much more quantitative results to be obtained in the near future, and other samples, particularly the supercalcine, will also be examined.

About 6% of the total caesium had undergone transmutations at the time of the most recent examinations. In the high caesium glasses, this represents nearly twice as many Cs $\rightarrow$ Ba transmutations as would occur in most commercial waste glasses. The other three glasses have experienced 25–50% of the transmutations as would a typical waste glass. But no changes in properties were found that would be of any consequence from a waste management viewpoint. The crystalline materials were also

relatively unaffected. However, the number of transmutations in these materials has been only about 13% of those expected for an actual waste.

This work was supported by the US Department of Energy contract DE-AC06-76RLO 1830.

Received 19 August 1981; accepted 17 February 1982.

1. Mendel, J. E. *et al.* *A State-of-the-Art Review of Materials Properties of Nuclear Waste Forms* PNL-3802, Pacific Northwest Laboratory, Richland (1981).
2. Roberts, F. P., Turcotte, R. P. & Weber, W. J. *Materials Characterization Center Workshop on the Irradiation Effects in Nuclear Waste Forms* (PNL-3588, Pacific Northwest Laboratory, Richland, 1981).
3. Permar, P. H. & McDonnell, W. R. *Proc. 10th int. Symp. on Effects of Radiation on Materials* (eds Kramer, D., Bragen, H. R. & Perrin, J. S.) 115–141 (ASTM-04-725000-35, 1981).

Synthesis of a quasi-stable kaolinite and heat capacity of interlayer water

P. M. Costanzo*, R. F. Giese Jr*, M. Lipsicas† & C. Straley†

*Department of Geological Sciences, State University of New York, 4240 Ridge Lea Road, Amherst, New York 14226, USA

† Schlumberger-Doll Research, Ridgefield, Connecticut 06877, USA

The mineral halloysite, in its hydrated and dehydrated forms, has attracted much attention¹ but, in spite of much careful X-ray and electron diffraction work, and electron microscopy, NMR and IR spectroscopic studies, there remain many questions about the crystal structure of the layers, their stacking, as well as the position, motion and bonding of the water molecules (in the hydrated form) to each other and to the surrounding silicate layers. The difficulties associated with studying halloysites are principally due to the great variability observed in natural samples² and, in the hydrated form, the existence of substantial amounts of pore water in addition to interlayer water. We have attempted to simplify the former difficulty by synthesizing a 10-Å hydrate similar to halloysite-10 Å (ref. 3) starting with a well characterized, crystalline kaolinite from Georgia. Here we report the successful synthesis of a 10-Å hydrate which is stable at reduced temperatures and contains no pore water. With this material we have measured C_p for the interlayer water between 110 and 270 K.

The question of the nature of the water in fully hydrated halloysite (designated here as halloysite-10 Å) is of great interest as it can serve as a model for water absorbed on silicate surfaces. In principle, it should be relatively simple to determine the chemical and physical properties of this water because there are few, if any, cations present and there is a much smaller net electrostatic charge on the clay layers than is found among all other hydrated layer silicate minerals. Natural hydrated halloysites must be kept in contact with liquid water or they rapidly dehydrate. The existence of large amounts of pore water makes it very difficult to study the properties of the interlayer water. Our approach has been to synthesize a dihydrate ($\text{Al}_2\text{Si}_2\text{O}_5(\text{OH},\text{F})_4 \cdot 2\text{H}_2\text{O}$) starting with well characterized kaolinites. This guarantees a reproducible material of known morphology, crystallinity and purity. The new synthesis produces a more stable product which is virtually free of pore water and thus represents a considerably simpler material from the point of view of spectroscopic and thermal analysis measurements.

We have previously described³ a means of producing a 10-Å hydrated kaolinite which is indistinguishable from natural halloysite-10 Å samples from Utah. As with the natural mineral,

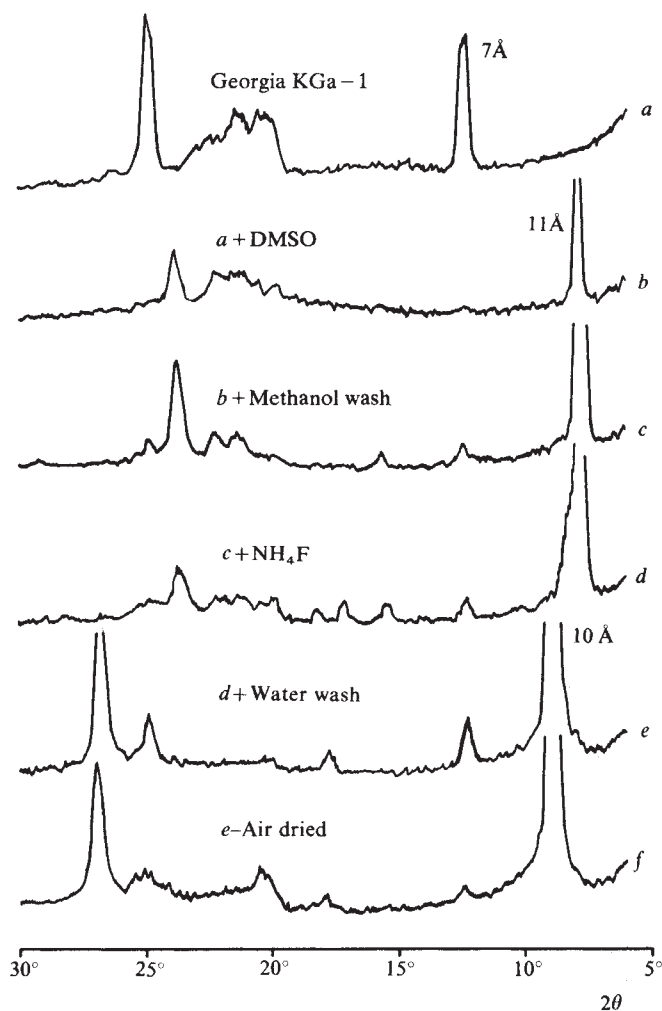


Fig. 1 X-ray diffractograms showing the original kaolinite (*a*), the intermediate stages of the synthesis (*b-d*) and the 10-Å product when wet (*e*) and dried (*f*). The diffractograms were made with CuK radiation.

this synthetic hydrate must be kept in contact with water to prevent spontaneous dehydration. In spectroscopic studies, such as IR and NMR, it is difficult, if not impossible, to distinguish between the pore water and the intercalated water.

The new synthesis is a variant of the method reported earlier. Figure 1 shows the sequence of steps in terms of the X-ray diffractograms of the starting material, the intermediate stages and the final dried product. The starting clay is the Georgia well crystallized kaolinite identified as KGa-1 (ref. 4). The clay is expanded with dimethyl sulphoxide (DMSO) containing 8% water (step *b*, Fig. 1) after which the intercalate is washed with methanol (step *c*, Fig. 1). The key step, as before, is the treatment of the intercalate (step *d*, Fig. 1) with ammonium fluoride (400 mg per g of kaolinite) for a period of several hours at 60 °C. This is followed by repeated water washings to remove the ammonium fluoride and intercalated DMSO. The wet clay is allowed to dry in ambient conditions. Complete loss of pore water may take several days. As reported earlier³, there is probably fluorine replacement of some surface hydroxyls. We are currently determining the extent of this substitution and its influence on hydrate formation.

Thermogravimetric analysis studies indicate a weight loss of 10% between 300 and 500 K. The theoretical weight loss for a dihydrate is 12.2%. Because X-ray diffraction indicates a yield of ~90%, the observed weight loss is in good agreement with the expected value. The remaining 10% is clay which has collapsed back to kaolinite. Moreover, thermogravimetric

analysis does not indicate any substantial amount of pore water. The synthetic hydrate is a hard, dry, massive material which will remain fully hydrated for periods of several months if kept in a sealed container at temperatures near 0 °C. It is not stable at room temperature and in fact loses water rapidly beginning at 280 K. Because of its instability at room temperature, we refer to the synthetic material as a quasi-stable kaolinite dihydrate.

To characterize the interlayer water further, we have studied the IR spectrum and taken measurements of the heat capacity (C_p). A more detailed discussion of these and other experiments (NMR and electron paramagnetic resonance) will be reported elsewhere.

The IR spectra (Fig. 2) were recorded on a Pye-Unicam 3-300 spectrometer with the clay dispersed in a fluorolube mull. The major features of the stretching bands of water and hydroxyls as well as the deformation bands of water are shown along with the corresponding spectra for the starting clay and for a natural halloysite-10 Å. The IR spectrum of the quasi-stable kaolinite dihydrate is clearly very similar to that of the natural 10-Å material. The greater crystallinity of the Georgia kaolinite starting material and the absence of pore water allow us to discern details in the spectrum of the synthetic hydrate which are difficult to see in the natural material. The broad band centred around 3,400 cm^{-1} is due to the continuous water layer⁵. In addition, one sees bands at 3,536 and 3,586 cm^{-1} which have been reported before⁵ but are very clearly resolved in our material. As will be discussed in detail elsewhere, these bands correspond to water molecules attached to di-trigonal sites on the tetrahedral surface of the interlayer space; we refer to these molecules as 'di-trigonal hole' water. The identification of the nature of the di-trigonal hole water has been facilitated by the existence of a kaolinite monohydrate ($d(001) = 8.4 \text{ Å}$) which is the dehydration product of the quasi-stable kaolinite-10 Å. Further evidence for the two types of water (di-trigonal hole and continuous monolayer) can be seen in the partially resolved doublet corresponding to the bending mode.

As noted by Cruz *et al.*⁶, there are only small shifts in the structural OH stretching frequencies between kaolinite and halloysite-10 Å. The same observation is true for the quasi-stable kaolinite dihydrate and its parent kaolinite (Fig. 2). This suggests that the surface hydroxyls are bonded to water molecules in much the same manner that they bond to the oxygens of the adjacent kaolinite layer.

Cruz *et al.*⁶ reported heat capacity (C_p) measurements of the water in a natural endellite (halloysite-10 Å). Our C_p measure-

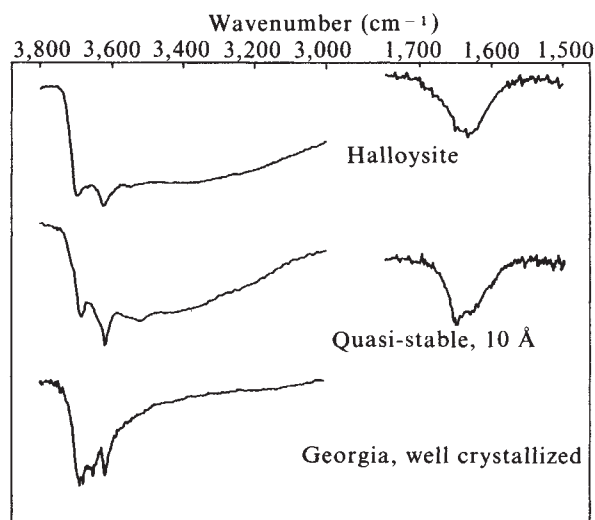


Fig. 2 IR spectra of the Georgia kaolinite (CMS-KGa), the synthetic 10 Å hydrate and a naturally occurring hydrated halloysite (Utah). The OH stretching and the water bending regions are shown.

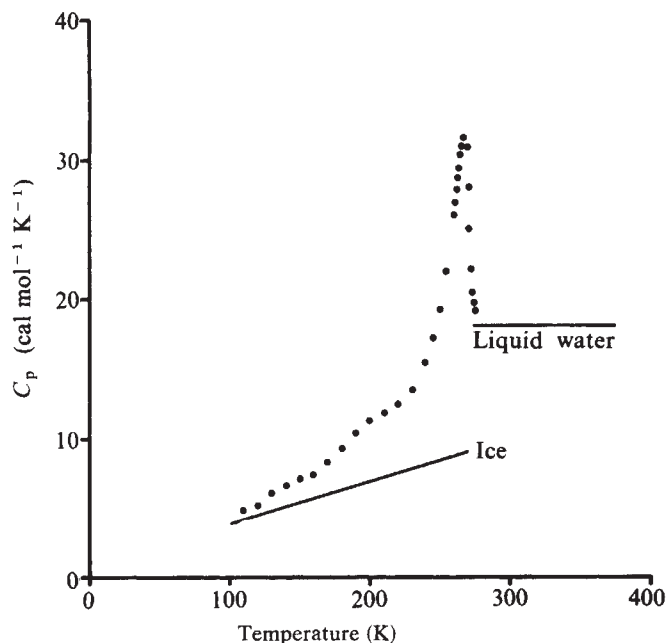


Fig. 3 A plot of heat capacity versus temperature for the inter-layer water of the quasi-stable kaolinite 10 Å.

ments were made on four samples (two between 110 and 190 K and two between 190 and 275 K) of the quasi-stable kaolinite-10 Å using a Perkin-Elmer DSC-2 scanning calorimeter equipped for low temperature operation. The experiment technique was that used by Cruz *et al.*⁶. The C_p values are plotted in Fig. 3. Internal agreement between different samples is of the order of 2–5% except for the very high C_p values in the vicinity of 265 K where the discrepancy is nearer 10%.

At the lowest temperatures, C_p values for the 10-Å water lie very close to values found for ice⁷. Beginning at ~160 K the C_p values move away from the ice line with a break in slope in the vicinity of 240 K. Above 240 K, C_p rises abruptly to a maximum at 260 K. This peak is reminiscent of the observations of Mraw and Naas-O'Rourke⁸ on 'freezable' pore water in coal. Cruz *et al.* did not report this peak in endellite because the upper temperature of their measurements was 255 K. By contrast, C_p for di-trigonal hole water alone, as measured in the kaolinite (8.4 Å) monohydrate, lies very close to the ice line throughout the temperature range studied but shows no discontinuity at 273 K. IR and NMR (T_1 spin-lattice relaxation time) measurements as a function of temperature suggest that the heat capacity peak of the kaolinite-10 Å water is associated with an 'order-disorder' type transition in which hole water is transferred to the continuous water monolayers as the temperature is raised. Note that one can readily distinguish two populations of interlayer water in the quasi-stable kaolinite 10 Å.

We thank Dr R. Prost, Dr M. Cruz-Fripiat and Professor J. J. Fripiat for their comments and suggestions concerning our investigation of water in kaolin minerals. Part of the research was supported by a grant from NSF/Sciences to R.F.G.

Received 23 November 1981; accepted 5 February 1982.

Volatiles in phyllosilicate minerals

David I. Norman & J. Michael Palin

New Mexico Institute of Mining and Technology, Socorro,
New Mexico 87801, USA

Volatiles other than water are known to occur in hydrous ring-structure silicates like beryl and cordierite^{1,2}. However, there has been limited study of volatiles in phyllosilicates. Ammonia has been reported in micas^{3,4}; Banks⁵ measured sulphur in biotite, sericite and chlorite by electron microprobe; and Norman⁶ reported evolution of H₂S and CO₂ when chlorite, sericite and kaolinite were heated to 500 °C. We have studied the thermal evolution of volatiles from phyllosilicates by mass spectrometry. The gases observed were, in approximate order of their abundance, CO₂, H₂S, H₂, SO₂, NH₃, CO, CH₄, N₂, Ar and He. These gases represent 0.35–9 wt % of the volatiles observed; the remainder is water. Step-heating studies indicate H₂S and CO₂ are not generated by the thermal breakdown of sulphide and carbonate minerals. Rather, they are bound within the phyllosilicate structures in a complex way. We show here that gases in chlorite, muscovite and biotite are related to their mode of origin. Phyllosilicates from hydrothermal ore deposits evolve several times more volatiles, principally CO₂ and H₂S, than phyllosilicates not from ore deposits.

Mineral separates were hand-picked, ground to -70+160 mesh, rinsed in distilled water, separated by heavy liquids, treated with warm 10% HCl for 15 min, multiply rinsed and oven dried. Samples were examined for impurities by scanning electron microscope (SEM) with energy dispersive detector, and loaded into a 5 × 40 mm quartz tube which was placed in an externally heated quartz furnace connected to a vacuum system. The sample was baked at 125 °C while being evacuated to 10⁻⁶ torr. Decrepitation was by heating to 700 °C for 1 h, then 1,100 °C for 1 h. Gases were determined by pressure measurement and quadrupole mass spectrometer; water was measured by weighing. Step-heating samples were heated for 30 min at 100 °C intervals and the evolved volatiles were analysed after each step.

Volatiles evolved on heating belong to five groups, water (including H₂), carbon gases, sulphur gases, nitrogen gases and rare gases. The C–S–N gases vary from 0.01 to 0.33 wt % of the phyllosilicates and analysed and 0.27–9.0 wt % of the volatiles released (Table 1). Within each group of gases, the ratio of oxidized to reduced species depends on the heating procedures. Step-heated samples yield more reduced gases than those heated in two steps.

The gases observed are related to the mode of origin of the mineral. Phyllosilicates from ore deposits evolve several times more gases than those not from ore deposits. In particular, chlorite associated with ore deposits evolves up to 10 times the CO₂ as other chlorite, and micas from ore deposits evolve up to 50 times more H₂S than micas from other environments.

Of particular concern was that the carbon and sulphur gases measured resulted from thermal breakdown of sulphide and carbonate mineral inclusions. Contamination of the mineral separates by sulphide and carbonate minerals is not considered to be serious because contaminate minerals, particularly sulphides, were readily detected by SEM examination and several chlorite samples from rock with a high percentage of sulphide minerals yielded little sulphur gases.

To investigate mineral inclusions and contaminate minerals as sources of carbon and sulphur gases, step-heating studies were done. Evolution of H₂S and CO₂ from phyllosilicate minerals is complex, similar to water (Fig. 1). Hydrogen sulphide is the principal sulphur gas with SO₂ evolution occurring primarily at $T \geq 1,000$ °C.

- Berthier, P. *Annls Chim. Phys.* **32**, 332–335 (1826).
- Sudo, T. *et al. Developments in Sedimentology* Vol. 31 (Elsevier, New York, 1981).
- Costanzo, P., Clemency, C. V. & Giese, R. F. *Clays Clay Miner.* **28**, 155–156 (1980).
- van Olphen, H. & Fripiat, J. J. *Data Handbook for Clay Materials and other Non-Metallic Minerals* (Pergamon, New York, 1979).
- Tarasevich, Yu I. & Gribna, I. A. *Kolloid. Zh.* **34**, 405–411 (1972).
- Cruz, M. I., Letellier, M. & Fripiat, J. J. *J. chem. Phys.* **69**, 2018–2027 (1978).
- Giauque, W. F. & Stout, J. W. *J. Am. chem. Soc.* **58**, 1144–1150 (1936).
- Mraw, S. C. & Naas-O'Rourke, D. F. *Science* **205**, 901–902 (1979).

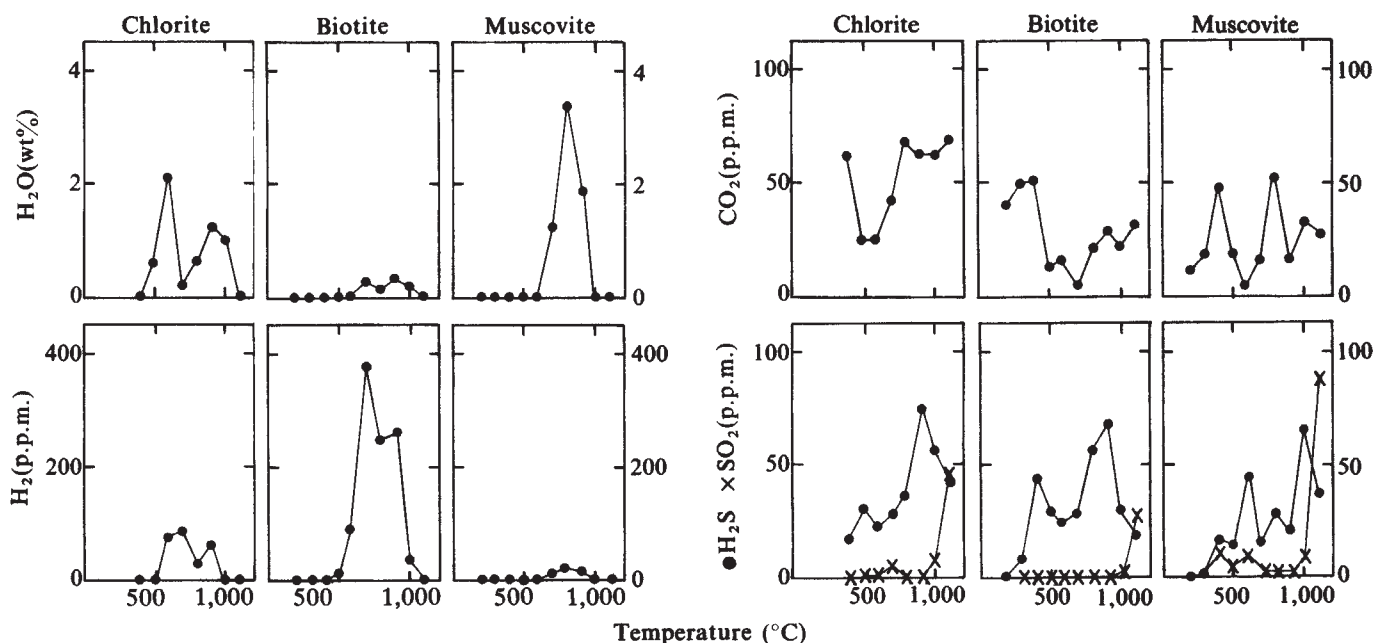


Fig. 1 Evolution of the principal volatiles when three different phyllosilicates were step heated. Chlorite is sample CHL2A from a chlorite schist. Muscovite is sample MUS2B from the Bjertnes pegmatite, Norway. Biotite is sample BIO2A from a pegmatite near Mora, New Mexico, USA.

In contrast, when ground sulphides were heated with a low-sulphur chlorite the thermal release curves are simple, SO_2 is the predominant sulphur gas, and considerable SO_2 is generated at temperatures $<1,000^\circ\text{C}$ (Fig. 2). When other mixtures of sulphides, which included FeS , FeS_2 , CuFeS_2 and PbS , were heated with chlorite similar results were obtained. Carbon dioxide generated by the thermal breakdown of carbonates similarly has a single thermal release curve which peaks at 700°C (Fig. 2).

We conclude that the sulphur and carbon gases released on heating phyllosilicates are not the result of mineral inclusions

or contaminate minerals because: (1) The thermal release of sulphur gases and CO_2 from phyllosilicates does not resemble the generation of these gases in our experiments. (2) Thermal release curves for H_2S and CO_2 are complex, with several peaks, and are similar to water which is structurally bound in phyllosilicates. (3) Contaminate sulphides evolve more SO_2 when heated with phyllosilicates and the SO_2 is generated at $T < 1,000^\circ$; whereas phyllosilicates evolve mostly H_2S and SO_2 is formed at $T > 1,000^\circ$. (4) Carbon dioxide thermal-release curves have a minimum at the temperatures at which carbonates breakdown; $700\text{--}800^\circ\text{C}$. (5) Gases bound to surfaces would be

Table 1 Volatiles released when phyllosilicates were heated to $1,100^\circ\text{C}$

Sample no. and location	H_2O (wt %)	H_2	CO_2	CO	CH_4	C_2H_6 (p.p.m.)	H_2S	SO_2	N_2	NH_3	Ar (p.p.b.)	He (p.p.b.)	(excluding H_2) wt % gases	$\frac{\text{gas}}{\text{H}_2\text{O}}$ $\times 100$
Chlorite														
PM-2A Pecos, NM USA (o)	7.50	—	2,200	240	34	—	15	67	63	—	—	NA	0.26	3.5
14-TBG-74 Tribag, Ontario, Canada (o)	10.1	—	1,800	140	15	—	31	2.7	—	—	—	NA	0.20	2.0
CHL4A Buchans, NFCanada (o)	11.1	20	1,500	150	230	—	52	25	36	—	—	—	0.20	1.8
CHL3A chlorite schist no. NM USA	10.7	42	380	17	2.9	—	490	240	—	28	—	NA	0.12	1.1
CHL3A (s)	10.1	260	420	14	4.2	—	310	76	—	51	—	NA	0.09	0.87
CHL2F meta-komatiite, NM USA	11.0	13	220	3.9	59	40	7.9	7.1	9.1	—	—	—	0.03	0.27
Muscovite														
CSB Creede, CO USA (o)	4.84	100	870	1,100	83	—	1,100	130	—	—	—	NA	0.33	6.8
MUS2B Bjertnes pegmatite, Norway	6.90	5.2	530	190	110	—	130	170	—	—	—	NA	0.11	1.6
MUS2B (s)	6.70	82	250	130	93	380	240	130	—	—	—	NA	0.12	1.8
MUS4A Sericite schist, no. NM, USA	0.76	8.8	86	5.5	7.4	25	2.2	0.6	7.5	—	—	4	0.01	1.3
MUS3A Pegmatite, NH USA	7.92	20	110	190	28	30	—	6.0	—	—	400	20	0.04	0.51
Biotite														
B101B Questa, NM USA (o)	2.70	110	170	490	26	—	1,000	500	31	226	—	NA	0.24	9.0
SR1A Santa Rita, NM USA (o)	2.90	340	180	14	2.0	—	880	65	—	10	—	NA	0.12	4.1
SR1B Santa Rita, NM USA (o)	1.60	47	730	220	140	82	160	44	27	—	1.2	0.4	0.14	8.8
CSG1A Cold Spring granite, MN, USA	2.80	320	210	4.5	44	24	170	58	14	5.9	72	—	0.05	1.8
MD1A Mayo-Darle Sn-granite Cameroon	7.96	55	110	—	45	27	17	5.2	23	—	32	—	0.05	0.63
B102A Pegmatite, no. NM USA	1.21	1,100	280	140	9.0	—	310	26	—	14	—	NA	0.08	6.6

Analyses of step-heated samples (s) and samples from ore deposits (o). Data are presented as the weight ratio of volatile to mineral. NA, not analysed.

released at $T < 500^\circ\text{C}$ and hence cannot account for most of the gases observed.

It is unlikely that volatiles other than rare gases occur in the phyllosilicates in the form they are measured. For example, water occurs as the hydroxyl ion. Neither the species which yield the gases observed nor their structural location can be deduced from the data presented or from calculated activation energies. There is evidence that nitrogen occurs as the NH_4 ion bound between silicate layers^{4,7}. Carbon dioxide production has been demonstrated by heating silicate minerals containing elemental carbon⁸; however, there is no reason to exclude the possibility that carbonate or bicarbonate ions might occur in phyllosilicates in a manner similar to ammonia. Our data suggest that water and hydrogen are released from chlorite and biotite at two temperatures by the breakdown of Fe-OH and Al-OH bonds; whereas these volatiles are evolved from muscovite primarily at one temperature by the breakdown of Al-OH bonds. Evolution of H_2S and CO_2 is not analogous to water in that they have multiple-peaked, thermal-release curves for muscovite. This suggests that H_2S and CO_2 are not structurally bound in the same way as water, although the temperature of gas release suggests bond energies for H_2S and CO_2 the same as the hydroxyl ion.

The association of high amounts of volatiles in phyllosilicates with ore deposits suggests that volatiles in phyllosilicates have

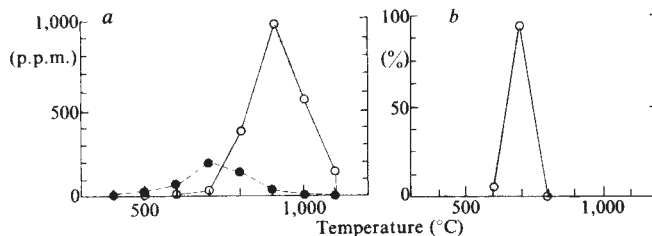


Fig. 2 *a*, An example of step-heating chlorite to which finely ground sulphides were added. $\circ$, SO_2 ; $\bullet$, H_2S . These data are for chlorite sample CHL2F to which was added ~ 0.35 wt % of a mixture of pyrite and chalcopyrite. *b*, Evolution of CO_2 ($\circ$) from finely ground calcite when it was step-heated in an analogous manner to the phyllosilicates.

a potential use as an exploration tool for ore mineralization. The ubiquitous presence of sulphur in phyllosilicates may be an unrecognized source of reduced sulphur. This reduced sulphur could be an important source of sulphur in the genesis of some ore deposits. Also, phyllosilicates, either during formation or through ion exchange processes, could be important in the inorganic reduction of sulphate.

This work was supported by Office of Surface Mining, US Department of the Interior grant no. G5104024.

Received 18 December 1981; accepted 21 January 1982.

1. Damon, P. E. & Kulp, J. L. *Am. Miner.* **43**, 433-459 (1958).
2. Beltrame, R. J., Norman, D. I. & Alexander, E. C. *EOS* **57**, 352 (1976).
3. Vedder, W. *Geochim. cosmochim. Acta* **44**, 221-228 (1965).
4. Honma, H. & Itihara, Y. *Geochim. cosmochim. Acta* **45**, 983-988 (1981).

5. Banks, N. G. *Econ. Geol.* **68**, 697-702 (1973).
6. Norman, D. I. thesis, Univ. Minnesota (1977).
7. Karyakin, A. V., Volynets, V. F., Kriventsova, G. A. & Vernadskiy, V. I. *Trans. Geochim. Inter.* **10**, 326-329 (1973).
8. Freund, F., Kathrein, H., Wengeler, H. & Knobel, R. *Geochim. cosmochim. Acta* **44**, 1319-1333 (1980).

Seismicity and rheology of subducted slabs

Rinus Wortel

Vening Meinesz Laboratory, Instituut voor Aardwetenschappen, University of Utrecht, Budapestlaan 4, 3584 CD Utrecht, The Netherlands

Since the incorporation of the inclined zones of intermediate and deep seismic activity into the theory of plate tectonics^{1,2} it has been generally accepted that earthquakes with focal depths ≥ 100 km occur in subducted slabs of (predominantly oceanic) lithosphere. Yet our understanding of the processes which generate intermediate and, in particular, deep earthquakes is still incomplete. This is largely due to our limited knowledge of the rheological behaviour of subducted oceanic lithosphere at depths between ~ 100 and 700 km. However, recently considerable insight has been gained into the rheology of oceanic lithosphere in near-surface conditions and I investigate here whether the seismic activity at depths > 100 km can be understood on the basis of this information. I first present results of temperature calculations performed for all subduction zones for which adequate data on seismicity, relative plate motion and age of the descending oceanic lithosphere exist. Subsequently, I adapt the temperature conditions governing the rheology near the surface by including depth-dependence. This leads to a depth-dependent critical temperature $T_c(z)$ above which the subducted lithosphere cannot sustain the stresses necessary to generate seismic events. Seismicity data and calculated temperature versus depth relationships are in good agreement with the rheological model predictions. This implies that the absence of seismic activity at depths ≥ 700 km should not be interpreted as direct evidence for the hypothesis that the 650-km discontinuity in the mantle acts as a barrier to vertical motion of subducted slabs.

McKenzie³ and Griggs⁴ investigated whether the maximum depth of earthquakes in subducted slabs is determined by a critical temperature. Although their results lent qualitative sup-

port to such a relationship their studies were based on limited data and hampered by great uncertainties concerning important geophysical parameters. Vlaar and Wortel⁵ have shown, from observational data, that the maximum depth of earthquakes in subduction zones increases with increasing age of the down-going oceanic lithosphere. Meanwhile, results of modelling the thermal structure of oceanic lithosphere^{6,7}, description of global relative plate motion⁸ and the widespread identification of magnetic anomalies in the oceans⁹ have provided the basis for more extensive and reliable calculations of the thermal structure of subducted slabs. The calculations reported here are based on McKenzie's approach³ in which a slab is assumed to sink in a mantle with an adiabatic temperature gradient. The adiabatic temperature distribution in the (upper) mantle below the lithosphere corresponds with a constant potential temperature $T_m^{(p)}$ (see ref. 3), for which we take $T_m^{(p)} = 1,200^\circ\text{C}$. Crough's⁷ model for the thermal evolution of oceanic lithosphere, with a temperature of $1,200^\circ\text{C}$ at the base of the lithosphere and fitted to bathymetry and heatflow data⁶, was used to calculate the age-dependent initial temperature distribution and thickness of the downgoing slab. This model implies an increase in lithospheric thickness from near-zero at a spreading centre to 101 km for 100 Myr old lithosphere. Apart from conduction and adiabatic compression the latent heat associated with the olivine-spinel phase change¹⁰, resulting in an increase in temperature of 135°C , was taken into account.

Several aspects of McKenzie's³ analytical model were critically examined by using a finite difference version of the heat conduction equation. Among these was the use of a constant thermal conductivity and the assumption of stationary input of the subduction process (constant lithospheric age at the trench). Model calculations with temperature-dependent thermal parameters as given by Schatz and Simmons¹¹ seemed to yield solutions which very closely agree with those for a constant diffusivity $\kappa = 0.9 \times 10^{-6} \text{ m}^2 \text{ s}^{-1}$. The maximum difference in minimum temperature (in $^\circ\text{C}$) was $\sim 2\%$, the calculations with temperature-dependent parameters giving the lower values. The constant value $\kappa = 0.9 \times 10^{-6} \text{ m}^2 \text{ s}^{-1}$ was used in the final calculations. Several subduction zones exhibit temporal variations in the age of the downbending lithosphere at the trench,

Table 1 Subduction zones and parameters

Zone	Age (Myr)	$v_c^{(n)*}$ (cm yr ⁻¹)	Length (km)	Dip (deg)	Depth (km)	Ref. [†]
1 Kuriles	100	7.9	880	50	660	26
2 Northern Honshu	125 → 100 [‡]	9.7	1300	30	605	2
3 Izu/Bonin	145 → 125	5.6 [§]	830	50–70	580	33
4 Ryukyu	60	5.6 [§]	450	45	275	33
5 Sumatra	70 → 60	5.8	400	35	200	34
6 Java	140 → 115	6.9	860	75	650	35
7 Tonga (20 °S)	100–140	8.9	960	48	680	36
8 Tonga (24 °S)	100–140	8.1	865	53	640	36
9 Kermadec (28 °S)	100–140	7.0	970	50	660	36
10 New Zealand	100–140	4.7	585	60	370	37
11 New Zealand	100–140	4.4	490	55	290	37
12 Chile (36–42 °S)	35	10.6	495	20	180	38
13 Chile (27–33 °S)	48 → 60	10.4	760	10	180	38
14 Chile (18–26 °S)	53 → 70	10.1	680	30	320	38
15 North-central Peru	45 → 65	9.1	750	10	185	38
16 Central America (87 °W)	40	8.4	275	65	200	39
17 Central America (90 °W)	40	7.6	355	60	250	40
18 Alaska (150 °W)	40 → 35	6.0	240 [¶]	≤30	172	41, 42
19 Aleutians (180 °W)	63 → 50	4.8	360	60	240	43

* Component of convergence rate normal to plate contact.

† References from which seismicity data (downdip length, dip and maximum depth) were taken. Length is measured from the trench to the depth of the deepest hypocentres. The dip angle is measured in the intermediate depth range.

‡ $n_1 \rightarrow n_2$ means that the present age at the trench is n_1 Myr and the deepest seismically active part of the slab was n_2 Myr old at the time it started to descend. n_2 was used in the thermal calculations.

§ From ref. 32.

|| Zones 10 and 11 correspond with Hamilton and Gale's³⁷ sections B and C respectively.

¶ Including a correction of 200 km for the excessively long accretionary prism.

for example, Sumatra, the Aleutians and South America (see ref. 12). Investigating the effect of these variations showed that the temperature distribution at a certain downdip distance in the subducted slab depends primarily on two parameters of the part of the lithosphere involved—the temperature distribution, or age, at the time of initial downbending at the trench and the time elapsed since it started to descend. Whether adjacent parts of the slab were slightly older or younger at their times of onset of subduction is negligible.

All subduction zones for which adequate data on slab geometry (length, dip and maximum depth of seismic zone), relevant plate kinematics and age of the downgoing slab were available are listed in Table 1. The main reason for rejecting subduction zones was lack of knowledge of real convergence rates, usually because of the unknown contribution of back-arc spreading. The relative plate motions given by Minster *et al.* and Jordan⁸ were used instead of the more recent results by Minster and Jordan¹³ because the former are more representative of the timespan which is involved in the subduction process. Age data were primarily taken from ref. 9. Seismicity data were taken from regional studies which report high-quality hypocentre locations (references are given in Table 1). See ref. 14 for a more extensive account of the data selection and the thermal modelling procedure.

In the present context the most relevant feature of the temperature distribution in a subducted slab is the minimum temperature at the depth of the deepest earthquake. For the subduction zones listed in Table 1 these temperatures are plotted in Fig. 1 as a function of depth. Owing to the nature of the thermal evolution of oceanic lithosphere the temperature distribution in a young descending slab is more sensitive to absolute errors in lithospheric age than the distribution in an old slab. For subducted slabs 35, 65 and 120 Myr old, an error of ± 10 Myr in age results in errors in minimum temperature at the deepest earthquake foci of $\pm 55^\circ\text{C}$, $\pm 25^\circ\text{C}$ and $\pm 15^\circ\text{C}$ respectively (apart from the latent heat contribution of the olivine–spinel phase change). The uncertainty in the age of the downgoing slab in zones 7–11 (see Table 1) results only in an uncertainty of $\pm 30^\circ\text{C}$ with respect to the values plotted in Fig. 1, which were calculated for an age of 120 Myr. Reliable calculations can only be made for continuous slabs. Therefore, the deep earthquakes beneath New Zealand (zone 11) and South America (zones 13, 14, 15), which probably occurred in

detached pieces of lithosphere, were not taken into account. In general, Fig. 1 displays a distinct increase in minimum temperature with depth. Molnar *et al.*¹⁵ have produced results which are qualitatively fairly similar to those in Fig. 1, but they did not concentrate on testing the conditions that govern the extent of a seismic zone.

The basis of the present discussion is that the rheology of the seismically active part of a slab must allow for the accumulation of stresses which are relaxed in the process of an

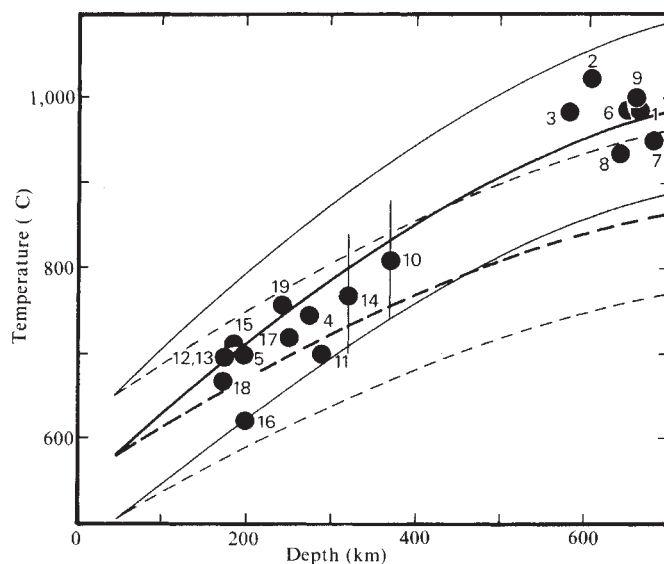


Fig. 1 Minimum temperatures at the depths of the deepest earthquakes in subduction zones: numbers refer to the zones listed in Table 1. The vertical bars for zones 10 and 14 reflect the uncertainty as to whether the olivine–spinel phase change has taken place. The points for zones 12 and 13 coincide. The thick solid curve represents T_{cr} (critical temperature above which no earthquakes are generated) as a function of depth, based on Stacey's²² mantle solidus. The thinner upper and lower solid curves indicate the range for T_{cr} resulting from uncertainties in the temperature at the depth of the 'calibration event' near the Tonga trench (see text). The thick dashed curve represents an alternative T_{cr} curve based on equation (10); the two other dashed curves indicate the uncertainty range for this T_{cr} curve.

earthquake. The problem is approached by relating the seismicity in the descending slab to relevant near-surface conditions and phenomena of which more direct observations are available. Two types of observation are considered: (1) observations of the geometry of the bending lithosphere near the trenches (in combination with those on seamount loading)^{16,17}, and (2) observations of the seismic activity of the oceanic lithosphere just before the latter starts to descend into the trench^{18,19}. Both data sets bear on the rheology of the lithosphere at the surface and, hence, on the input of the subduction process (section A in Fig. 2).

The rheological model used by Caldwell and Turcotte¹⁶ for the bending and seamount loading is simpler than that used by Bodine *et al.*¹⁷. In view of the many uncertainties attached to the slab's rheology at greater depths it seems more appropriate to base the analysis on Caldwell and Turcotte's¹⁶ model. These authors assumed that below a certain temperature T_e elastic stresses do not relax on geological time scales whereas at temperatures above T_e stresses are relaxed by creep processes. In this spirit Caldwell and Turcotte¹⁶ studied the rheology of oceanic lithosphere, using power-law creep behaviour for olivine appropriate for differential stresses below 2 kbar (ref. 20)

$$\frac{d\epsilon_p}{dt} = C\sigma^3 \exp(-A/RT) \quad (1)$$

$d\epsilon_p/dt$ is the creep rate, σ the differential stress, T absolute temperature, R the universal gas constant, A an activation volume and C a constant. By using a Maxwell-type of visco-elastic model they derived an expression for T_e :

$$T_e = A / \{R \ln (\frac{2}{3} EC\tau_R \sigma_0^2)\} \quad (2)$$

where E is Young's modulus and τ_R is the time for the stress σ_0 , applied at $t = 0$, to relax to one-half its original value. Using Goetze's²⁰ values $A = 122 \text{ kcal mol}^{-1}$ and $C = 70 \text{ bar}^{-3} \text{ s}^{-1}$, and $E = 6.5 \times 10^{10} \text{ N m}^{-2}$, for $0.1 < \sigma_0 < 2.0 \text{ kbar}$ and $0.1 < \tau_R < 10.0 \text{ Myr}$ T_e varies from 660 to 840 °C. These temperatures are in good agreement with those derived, from studies of lithospheric bending¹⁶, for the base of the elastic upper part of the lithosphere and also for the depth corresponding with Bodine *et al.*'s¹⁷ mechanical thickness H of the lithosphere (defined as the depth below which the yield strength of the lithosphere is $< 500 \text{ bar}$).

The usual way²¹ to include depth- or pressure-dependence in equation (1) is to put

$$\frac{d\epsilon_p}{dt} = C\sigma^3 \exp(-gT_m/T) \quad (3)$$

where g is a constant and T_m the melting temperature (in K) at the appropriate depth or pressure. The analogue of equation (2) is then

$$T_e = gT_m / \ln (\frac{2}{3} EC\tau_R \sigma_0^2) \quad (4)$$

For $0.1 < \tau_R < 10.0 \text{ Myr}$ and $0.1 < \sigma_0 < 2.0 \text{ kbar}$ we find

$$55.2 < \ln (\frac{2}{3} EC\tau_R \sigma_0^2) < 65.8 \quad (5)$$

As a good approximation, we may write instead of equation (4):

$$T_e \approx \alpha T_m \text{ with } \alpha = g/60 \quad (6)$$

Briefly, studies of the bending of oceanic lithosphere have shown that the rheology of the lithosphere is strongly temperature dependent. The temperature-dependence is controlled by the ratio T/T_m .

The second data set comprises precise depth determinations of earthquakes associated with the bending of the lithosphere near the trenches given by Chapple and Forsyth¹⁸ (see also ref. 19). From this data set we can derive an estimate of the maximum temperature at which deformation of the lithosphere (in near-surface conditions) produces earthquakes. This maximum temperature is found for the deepest event of the data set, which occurred east of the Tonga trench at a depth

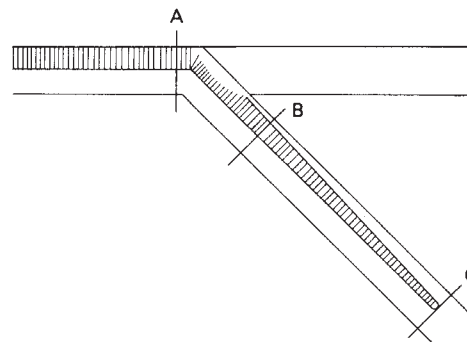


Fig. 2 Representation of a subducted slab. Hatching indicates the parts of the downgoing slab which are potentially seismically active. Earthquakes at or near the contact between the descending slab and the upper plate are not considered, in view of their different character. The lithosphere at section A is representative of the input of the subduction process. Section B and C are at the depth of the double seismic zones (about 120–200 km) and at the downdip end of the seismic zone respectively.

of $45 \pm 5 \text{ km}$. Its focal mechanism solution (horizontal compression) indicates that it took place below the neutral surface of the bending plate. The age of the lithosphere at the Tonga trench, which determines the temperature distribution, is estimated to be in the range 100–140 Myr. According to Crough's⁷ thermal model the temperature at the hypocentral depth of this event is $580 \pm 70 \text{ °C}$, the error being the r.m.s. error due to the uncertainties in lithospheric age and focal depth. This temperature is taken to be the (empirical) maximum temperature, designated as T_{cr} , at which the deformation of oceanic lithosphere is accompanied by seismic activity. For the analysis of intermediate and deep seismicity we are interested in the depth-dependence of T_{cr} . Because the compressive 'bending' earthquake occurred in the depth range of the lithosphere of which the rheology is controlled by temperature and on the basis of the results obtained for T_e (see equations (4)–(6)), we assume

$$T_{cr}(z) = \beta T_m(z) \quad (7)$$

where T_m is the melting temperature and β is a constant. For a chosen $T_m(z)$ curve β can be determined from the intercept-value, that is $T_{cr} = 580 \pm 70 \text{ °C}$ at the depth of the Tonga event ($\approx 45 \text{ km}$). For T_m we first take Stacey's²² composite mantle solidus T_s . For β we find $\beta = 0.54 \pm 0.04 \equiv \beta_s$. $T_{cr}(z)$ and the temperature range resulting from the uncertainty in β are plotted in Fig. 1. As pointed out by Anderson and Minster²³, the composite mantle solidus is representative of the low melting point phases in the polyphase mantle system. The effect of different choices for T_m can be illustrated by using, as an extreme alternative, the liquidus for which we take 2,163 K at the surface. This yields a lower value of β : $\beta = 0.38 \pm 0.03 \equiv \beta_l$. According to Kennedy and Higgins²⁴ the liquidus temperature of the mantle increases more rapidly with depth than the solidus temperature, although in the upper mantle down to 700 km this effect may not be very pronounced. Thus

$$\frac{dT_l}{dz} \geq \frac{dT_s}{dz} \quad (8)$$

But, because $\beta_l/\beta_s \approx 0.7$, we have

$$\beta_l \frac{dT_l}{dz} < \beta_s \frac{dT_s}{dz} \quad (9)$$

The upper bound of the temperature range for $T_{cr}(z)$, based on Stacey's mantle solidus, is taken as an upper bound for all T_{cr} curves to be considered. A lower bound can be constructed by taking

$$\frac{dT_{cr}}{dz} = \beta_l \frac{dT_s}{dz} \quad (10)$$

The T_{cr} curve based on equation (10) is also plotted in Fig. 1 (thick dashed curve). The two other dashed curves (upper and lower) indicate the range of T_{cr} resulting from the uncertainty in β_1 .

In the downbending lithosphere near the trench the minimum temperature is clearly below T_{cr} . From numerical model calculations I infer that, for relevant parameter values, the minimum temperature T_{min} in a subducted slab satisfies

$$\frac{dT_{min}(z)}{dz} > \frac{dT_{cr}(z)}{dz} \quad (11)$$

for both T_{cr} curves shown in Fig. 1. Because $T_{min}(z)$ depends on the initial temperature distribution (or age) of the downgoing slab, the direction and rate of plate convergence and the dip of the slab, the depth at which $T_{min}(z) = T_{cr}(z)$ varies from one zone to another. Whether the downdip extent of a seismic zone is determined by a limiting temperature T_{cr} can now be checked at the level of section C in Fig. 2: the minimum temperatures at the depths of the deepest foci in continuous subducted slabs, plotted as a function of depth, should delineate the T_{cr} curve. Figure 1 shows that all calculated minimum temperatures fall between the upper and lower bounds derived for T_{cr} . Even with the more restricted range belonging to Stacey's solidus curve the temperatures are in very good agreement. Note that taking into account the effect of stress level variations [see equation (5)] or errors in the melting point curve would broaden the model range for T_{cr} .

An additional check of the model (seismic activity confined to regions with $T < T_{cr}$) can be made at the level of section B in Fig. 2 (120–200 km depth), where detailed hypocentral location procedures have revealed a double seismic zone^{25–27}. The distance between the two parallel zones varies from ~25 km for the Aleutians²⁵ to 30–40 km for the Kuriles²⁶ and Honshu²⁷. The temperature distributions I have calculated for the subducted slabs in these regions show cold zones ($T < T_{cr}$) with thicknesses exceeding the above cited values by 5–10 km.

Therefore, I conclude that the gross features of subduction zone seismicity at depths between ~100 and 700 km are adequately explained by a model in which seismic activity is confined to those parts of the subducted slabs which have temperatures below a depth-dependent critical value T_{cr} .

For many years, deep seismic activity has influenced debates about the depth extent of mantle convection^{28–31}. From the results reported here I conclude that the absence of seismic activity at depths >700 km as such should not be considered as direct evidence for the hypothesis that subducted slabs cannot penetrate the 650-km mantle discontinuity.

On the basis of our knowledge of the rheology of oceanic lithosphere and melting temperatures in the mantle, the maximum depth of earthquakes can be understood in terms of thermal assimilation or rheological de-activation.

Arguments involving the depth variations of earthquake magnitudes or energy release have not been used here in order to avoid confusion; the slab's rheology as considered here must provide basic conditions for generation of earthquakes, whereas magnitudes and energy release also depend on the forces acting on the slab and the physical mechanism of an earthquake.

I thank Professor N. J. Vlaar and Dr S. Cloetingh for discussions and comments and Dr D. W. Forsyth for a preprint.

Received 24 November 1981; accepted 17 February 1982.

- Isacks, B., Oliver, J. & Sykes, L. R. *J. geophys. Res.* **73**, 5855–5899 (1968).
- Isacks, B. & Molnar, P. *Rev. Geophys. Space Phys.* **9**, 103–174 (1971).
- McKenzie, D. P. *Geophys. J. R. astr. Soc.* **18**, 1–32 (1969).
- McKenzie, D. P. *Tectonophysics* **10**, 357–366 (1970).
- Griggs, D. T. in *The Nature of the Solid Earth* (ed. Robertson, E. C.) 361–384 (McGraw-Hill, New York, 1972).
- Vlaar, N. J. & Wortel, M. J. R. *Tectonophysics* **32**, 331–351 (1976).
- Parsons, B. & Sclater, J. G. *J. geophys. Res.* **82**, 803–827 (1977).
- Crough, S. T. *Nature* **256**, 388–390 (1975).
- Minster, J. B., Jordan, T. H., Molnar, P. & Haines, E. *Geophys. J. R. astr. Soc.* **36**, 541–576 (1974).
- Jordan, T. H. *J. geophys. Res.* **80**, 4433–4439 (1975).
- Pitman, W. C. III, Larson, R. L. & Herron, E. M. *The Age of the Ocean Basins and Magnetic Lineations of the Ocean* (Charts, Geological Society of America, Boulder, 1974).
- Turcotte, D. L. & Schubert, G. *J. geophys. Res.* **76**, 7980–1987 (1971).

- Schatz, J. F. & Simmons, G. J. *Geophys. Res.* **77**, 6966–6983 (1972).
- Wortel, M. J. R. & Vlaar, N. J. *Phys. Earth planet. Inter.* **17**, 201–208 (1978).
- Minster, J. B. & Jordan, T. H. *J. geophys. Res.* **83**, 5331–5354 (1978).
- Wortel, R. thesis, Univ. Utrecht (1980).
- Molnar, P., Freedman, D. Shih, J. S. F. *Geophys. J. R. astr. Soc.* **56**, 41–54 (1979).
- Caldwell, J. G. & Turcotte, D. L. *J. geophys. Res.* **84**, 7572–7576 (1979).
- Bodine, J. H., Steckler, M. S. & Watts, A. B. *J. geophys. Res.* **86**, 3695–3707 (1981).
- Chapple, W. M. & Forsyth, D. W. *J. geophys. Res.* **84**, 6729–6749 (1979); Forsyth, D. W. *Phys. Earth planet. Inter.* (in the press).
- Chen, T. & Forsyth, D. W. *J. geophys. Res.* **83**, 4995–5003 (1978).
- Goetze, C. *Phil. Trans. R. Soc. London A* **288**, 99–119 (1978).
- Weertman, J. & Weertman, J. R. A. *Rev. Earth planet. Sci.* **3**, 293–315 (1975).
- Stacey, F. D. *Phys. Earth planet. Inter.* **15**, 341–348 (1977).
- Anderson, D. L. & Minster, J. B. *Prof. J. Coulomb Symp. on Source Mechanism and Earthquake Prediction* (1980).
- Kennedy, G. C. & Higgins, G. H. *Tectonophysics* **13**, 221–232 (1972).
- Engdahl, E. R. & Scholz, C. H. *Geophys. Res. Lett.* **4**, 473–476 (1977).
- Veith, K. F. thesis, Southern Methodist Univ., Dallas (1974).
- Hasegawa, A., Umino, N., Takagi, A. & Suzuki, Z. *Tectonophysics* **57**, 1–6 (1979).
- Davies, G. F. *Geophys. J. R. astr. Soc.* **49**, 459–486 (1977).
- O'Connell, R. J. *Tectonophysics* **38**, 119–136 (1977).
- Richter, F. M. *J. geophys. Res.* **84**, 6783–6795 (1979).
- Peltier, W. R. in *Physics of the Earth's Interior* (eds Dziewonski, A. M. & Boschi, E.) 361–431 (North Holland, Amsterdam, 1980).
- Seno, T. *Tectonophysics* **42**, 209–226 (1977).
- Katsumata, M. & Sykes, L. R. *J. geophys. Res.* **74**, 5923–5948 (1969).
- Hamilton, W. *Geol. Surv. profess. Pap.* **1078** (1979).
- Fitch, T. J. & Molnar, P. *J. geophys. Res.* **75**, 1431–1444 (1970).
- Sykes, L. R. *J. geophys. Res.* **71**, 2981–3006 (1966).
- Hamilton, R. M. & Gale, A. W. *J. geophys. Res.* **73**, 3859–3876 (1968).
- Barazangi, M. & Isacks, B. L. *Geology* **4**, 688–692 (1976).
- Dewey, J. W. & Algermissen, S. T. *Bull. seism. Soc. Am.* **64**, 1033–1048 (1974).
- Molnar, P. & Sykes, L. R. *Bull. geol. Soc. Am.* **80**, 1639–1684 (1969).
- Tobin, D. G. & Sykes, L. R. *J. geophys. Res.* **71**, 1659–1667 (1966).
- Jacob, K. H., Nakamura, K. & Davies, J. N. in *Island Arcs, Deep Sea Trenches and Back-arc Basins* (eds Talwani, M. & Pitman, W. C. III), 243–258 (American Geophysical Union, Washington DC, 1977).
- Engdahl, E. R. in *Island Arcs, Deep Sea Trenches and Back-arc Basins* (eds Talwani, M. & Pitman, W. C. III) 259–272 (American Geophysical Union, Washington DC, 1977).

Indonesian Permian brachiopod fauna and Gondwana–South-East Asia relationships

N. W. Archbold*, C. J. Pigram†, N. Ratman‡ & S. Hakim‡

* Department of Geology, University of Melbourne, Parkville, 3052, Victoria, Australia

† Bureau of Mineral Resources, Geology and Geophysics, Jalan Cilaki, 49, Bandung, Indonesia

‡ Geological Research and Development Centre, Bandung, Indonesia

The island of New Guinea, geologically and structurally part of the Indian–Australian plate^{1,2}, is shown on continental reconstructions of the Permian globe as forming the northeastern part of the supercontinent of Gondwana^{3,4}, facing a large Tethys Ocean to the north. The existence of an ocean separating South-East Asia from Gondwana by some 45° of latitude during the Permian has been widely disputed because of the stratigraphical, structural^{5–10} and palaeontological^{11–14} links between the regions. Palaeontological comparisons of New Guinea and Asia are essential for testing whether or not South-East Asia was a part of Gondwana during the Permian^{15,16}. We now give an account of the first, diverse, reliably dated, Permian articulate brachiopod fauna to be discovered from the island of New Guinea. This fauna is interpreted as reflecting the geographical proximity of Thailand and Irian Jaya during the late early Permian.

Reports of Permian marine faunas from Papua New Guinea^{17,18} have been shown to be erroneous¹⁹, the faunas being Triassic in age. Previous collections of late Palaeozoic marine faunas from Irian Jaya have been inadequate for precise dating and have provided only a generalized Permo-Carboniferous age¹. The new collections, from the Birds Head region of Western Irian Jaya, are from four localities within the Aifat Formation of the Aifat Group, outcropping on the Taminabuan 1:250,000 sheet area²⁰. The Aifat Formation is some 1,100 m thick and consists chiefly of dark calcareous shale and marl with limestone intercalations. Three of the localities have brachiopod faunas which are remarkably similar to those from the Rat Buri Limestone of Thailand^{20,21} of late

Baigendzhinian (latest Artinskian)²¹ or early Kungurian²² age. They also have some links with brachiopod faunas of Western Australia and Bituani, Timor. The fourth locality has yielded a fauna, sharing no species or genera in common with the other three localities, which contains a species of the rare chonetacean *Quinquenella* that is morphologically close to a Western Australian species²³ of the genus. This comparison is used to date the Irian Jaya species as late Baigendzhinian. The locality, in addition to *Quinquenella*, has only yielded fragmentary spiriferid brachiopods which are unidentifiable at this stage. Of the 14 genera from the Irian Jaya fauna, 13 are present in the Rat Buri Limestone faunas; 11 species are morphologically close between the two regions (Table 1).

Most of the classic Permian brachiopod faunas of Timor, such as the Basleo fauna, are much younger than the Irian Jaya and Thailand faunas²¹. An exception is the Bituani fauna which exhibits some similarities with the Irian Jaya fauna (such as representatives of *Stereochia* and *Stenosisma* close to *Stenosisma* sp. nov.) and the Thailand and Western Australian late early Permian faunas²¹ (such as the presence of *Retimarginifera*). However, the Bituani fauna has not been fully described thus preventing detailed comparison. Furthermore, the tectonic setting of the Bituani fauna is confused because the Timor limestones occur as lenses in the Tertiary Bobonaro Scaly Clay Olistostrome. They have been interpreted as being allochthonous tropical limestones from the northern margin of Tethys²⁴. However, limestone is known from the Australian north-west offshore shelf² and the Timor limestone faunas are subtropical, not tropical, and show greatest taxonomic affinities to the faunas of northern India and north-west Australia². The Timor limestones are probably of northern Gondwanan affinity². Links between the late early Permian brachiopod faunas of Thailand and Western Australia, although previously considered to be minimal²⁵, are becoming more apparent as the faunas of Western Australia become better known²⁶. Similarities between Western Australian and South-East Asian Permian molluscan faunas are also significant²⁷.

A Permian geographical position for Thailand close to that of Irian Jaya and hence to the northern margin of Gondwana is a possible explanation of these faunal links and the mixed Cathaysian–*Glossopteris* Permian flora from Thailand that has previously been explained by parallel evolution²⁸. The mixed Permian floras of Thailand and Irian Jaya¹ were apparently on the northern fringe of the Gondwana *Glossopteris* Floral Realm. The Irian Jaya *Glossopteris* occurs in stratigraphically concordant sediments above the Aifat Formation. The source of the Permian rock sequence of the Birds Head Region is most probably the large area of basement (Kemum Formation) exposed to the north of the outcrops of Permian rocks. The Kemum Formation of Silurian age, is an autochthonous

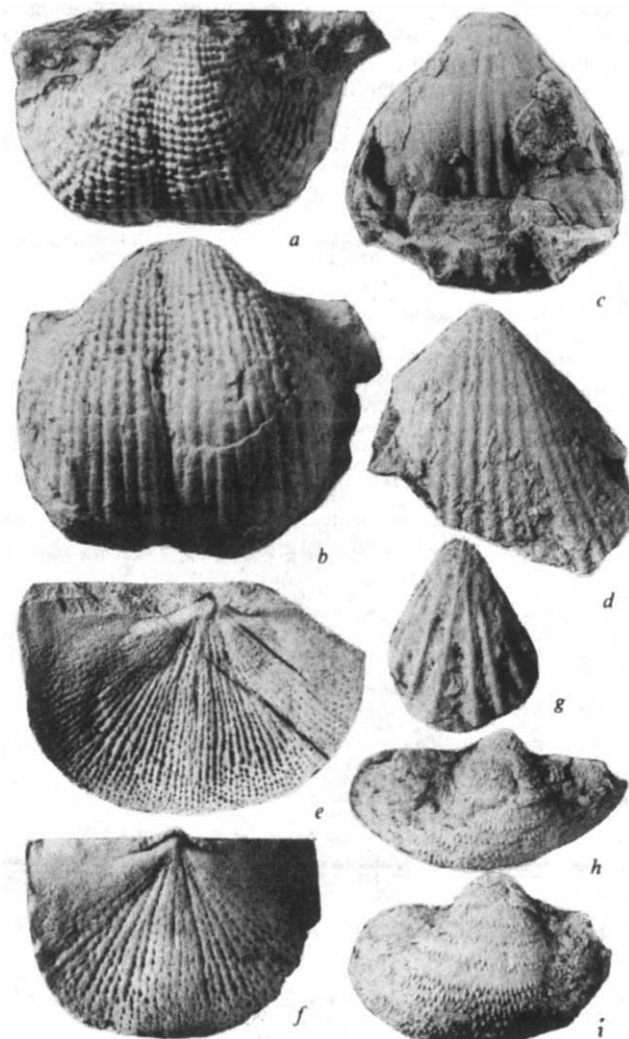


Fig. 1 Permian brachiopods from Irian Jaya. a, b, Ventral posterior and ventral views of *Stereochia* sp. nov., $\times 1$. c, Ventral view of *Stenosisma* sp. nov., $\times 2$. d, Ventral view of *Stenosisma* sp. cf. *S. tetricum* Grant 1976, $\times 2$; e and f are views of two internal moulds of dorsal valves of *Quinquenella* sp. nov., $\times 4.5$ and $\times 5$ respectively; g is a ventral view of *Hustedia* sp. cf. *H. ratburiensis* Waterhouse & Piyasin 1970, $\times 2$; h and i are ventral posterior and ventral views of *Stictozoster* sp. cf. *S. leptus* Grant 1976, $\times 2$.

Table 1 Late early Permian brachiopods from Irian Jaya²⁰ and a comparison with those from the Rat Buri Limestone²¹

Irian Jaya	Thailand
<i>Streptorhynchus</i> sp.	<i>Streptorhynchus khwaense</i>
<i>Rhipidomella</i>	Present*
<i>Chonetinella</i> sp. nov.	Present*
<i>Stictozoster</i> sp. cf. <i>S. leptus</i>	<i>Stictozoster leptus</i>
<i>Stereochia</i> sp. nov.	<i>Stereochia litostyla</i>
<i>Linoproductus</i> sp. nov.	<i>Linoproductus</i> sp. indet.
<i>Cancrinella</i> sp.	Present*
<i>Stenosisma</i> sp. nov.	<i>Stenosisma</i> sp. A
<i>Stenosisma</i> sp. cf. <i>S. tetricum</i>	<i>Stenosisma tetricum</i>
<i>Cruricella</i> sp.	<i>Cruricella couria</i>
<i>Callispirina</i> sp.	<i>Callispirina austrina</i>
<i>Spiriferellina</i> sp.	<i>Spiriferellina yanagidai</i>
<i>Hustedia</i> sp. cf. <i>H. ratburiensis</i>	<i>Hustedia ratburiensis</i>
<i>Cleiothyridina</i> sp.	<i>Cleiothyridina tribulosa</i>
<i>Quinquenella</i> sp. nov.	—

* The genus is present in the Thailand faunas but the species is dissimilar to the Irian Jaya species.

sequence within the Birds Head Region. A palaeomagnetic investigation of the Birds Head Region is being undertaken by the Bureau of Mineral Resources, Geology and Geophysics.

The new palaeontological evidence presented above supports the incorporation of much of South-East Asia onto the northern margin of Gondwana¹⁰. Permian oceanic Tethys appears to have been north of Southern Thailand, because systematic analysis of the brachiopod fauna has indicated ease of migration between Southern Thailand, Timor, Irian Jaya and Western Australia of closely related species in the late early Permian²⁰. Articulate brachiopods appear to have a free-swimming larval stage of a few hours or days²⁹. This limiting factor, in addition to latitudinal temperature zones and presumed surface currents of the Permian oceans³⁰, would have been a major obstacle to brachiopod migration across a vast Tethys ocean. While several Permian brachiopod genera do exhibit a disjunct distribution, occurring in both Gondwanan and Boreal faunas³¹, the similarity at the species level between the Irian Jaya and Thailand faunas apparently reflects the geographical proximity of the two regions. The similarity of species seems to imply a relatively narrow latitudinal spread for Irian Jaya and Thailand during the late early Permian, possibly within 5°–10°, because species may range widely along latitudes but they tend not to range far across latitudes³².

We thank Dr H. M. S. Hartono, the director, Geological Research and Development Centre, Bandung, Indonesia and the director, Bureau of Mineral Resources, Australia for permission to publish, and Dr S. K. Skwarko for continued advice.

Received 20 November 1981; accepted 5 February 1982.

1. Visser, W. A. & Hermes, J. J. *Verh. K. ned. geol.-mijnb. Genoot.* **20**, 1–265 (1962).
2. Hamilton, W. *Prof. Pap. U.S. geol. Surv.*, 1078, 1–345 (1979).
3. Briden, J. C., Drewry, G. E. & Smith, A. G. *J. Geol.* **82**, 555–574 (1974).
4. Ziegler, A. M., Scotese, C. R., McKerrow, W. S., Johnson, M. E. & Bambach, R. K. *A. Rev. Earth planet. Sci.* **7**, 473–502 (1979).
5. Ridd, M. F. *Nature* **234**, 531–533 (1971).
6. Acharyya, S. K. *Tectonophysics* **56**, 261–276 (1979).
7. Zonenschein, L. P. & Gorodnitskii, A. M. *Geotektonika* **3**, 3–24 (1977).
8. Audley Charles, M. G., Carter, D. J. & Milson, J. S. *Nature phys. Sci.* **239**, 35–39 (1972).
9. Sengor, A. M. C. *Nature* **279**, 590–593 (1979).
10. Mitchell, A. H. G. *J. geol. Soc. Lond.* **138**, 109–122 (1981).
11. Hart, G. F. *S.E.P.M. Spec. Publ.* **21**, 148–164 (1974).
12. Kremp, G. O. W. *Rev. Palaeobot. Palynol.* **17**, 113–132 (1974).
13. Termier, H. & Termier, G. C. *re heb. Séanc. Acad. Sci., Paris* **283D**, 139–142 (1976).
14. Ross, J. P. R. *Historical Biogeography, Plate Tectonics and the Changing Environment* (eds Gray, J. & Boucot, A. J.) 259–278 (Oregon State University Press, Corvallis, 1979).
15. Ridd, M. F. *Nature phys. Sci.* **240**, 140–141 (1972).
16. Ridd, M. F. *J. geol. Soc. Lond.* **137**, 635–640 (1980).
17. Glaessner, M. F., Llewellyn, K. M. & Stanley, G. A. V. *Austral. J. Sci.* **13**, 24–25 (1950).
18. Rickwood, F. K. *J. geol. Soc. Aust.* **2**, 63–82 (1955).
19. Skwarko, S. K., Nicoli, R. S. & Campbell, K. S. W. *B.M.R. J. Aust. Geol. Geophys.* **1**, 219–230 (1976).
20. Archbold, N. W. *Geol. Res. Dev. Cen. Bandung, Indonesia, Pal. Ser.* **2** (1) (in the press).
21. Grant, R. E. *J. Palaeont.* **50**, 1–269 (1976).
22. Waterhouse, J. B. *Pap. Dep. Geol. Univ. Qd* **7**, 1–230 (1976).
23. Archbold, N. W. *J. Paleont.* **55**, 204–210 (1981).
24. Barber, A. J. & Audley-Charles, M. G. *Tectonophysics* **30**, 119–128 (1976).
25. Griffiths, J. & Burrett, C. *Nature phys. Sci.* **245**, 92–93 (1973).
26. Archbold, N. W. *Proc. Roy. Soc. Vict.* **91**, 181–192 (1980).
27. Dickens, J. M. *Aust. nat. Hist.* **17**, 270–273 (1972).
28. Asama, K. *Geology and Palaeontology of South East Asia Vol. 2* (eds Kobayashi, T. & Toriyama, R.) 197–237 (University of Tokyo Press, 1966).
29. Rudwick, M. J. S. *Living and Fossil Brachiopods* (Hutchinson, London, 1970).
30. Valentine, J. W. & Moores, E. M. *Mem. Can. Soc. petrol. Geol.* **2**, 759–766 (1973).
31. Waterhouse, J. B. *Nature* **216**, 47–49 (1967).
32. Waterhouse, J. B. & Gupta, V. J. *J. geol. Soc. Ind.* **20**, 461–464 (1979).

Lateglacial and early Flandrian chronology of the Isle of Mull, Scotland

M. J. C. Walker

Department of Geography, Saint David's University College, Lampeter, Dyfed SA48 7ED, UK

J. J. Lowe

Geography Section, City of London Polytechnic, Calcutta House, Old Castle Street, London E1 7NT, UK

The Isle of Mull, which lies off the west coast of Scotland (Fig. 1), was partially over-ridden by ice from the mainland during the build-up of the late Devensian ice sheet¹ but, in common with the adjacent islands of Rum² and Skye³, supported smaller ice masses, including independent valley and cirque glaciers during the Loch Lomond Stadial^{4,5} (approximately equivalent with the Younger Dryas of Scandinavia^{6,7}). However, although the recent glacial history of Mull is understood, at least in general terms, little is known about the nature and timing of the major environmental changes that preceded and succeeded the Loch Lomond Stadial on the island. We present here pollen-stratigraphical evidence supported by 15 radiocarbon dates which provide new information on landscape changes on the Isle of Mull from the wastage of the last Scottish ice sheet to the early part of the Flandrian period. The radiocarbon dates are of particular significance as they form the basis for the first Lateglacial and early Flandrian chronology for the islands of the Scottish Inner Hebrides.

Four sites have been investigated (Fig. 1). Loch an t-Suidhe (grid ref. NM370215, altitude 30 m OD) on the Ross of Mull and Mishnish (grid ref. NM455565, altitude 120 m OD) to the west of Tobermory are infilled basins which lie outside the area

affected by Loch Lomond Stadial ice, and contain a typical Lateglacial sequence⁸ of organic muds between minerogenic sediments, these deposits in turn being overlain by Flandrian muds and peats. Coire Clachach (grid ref. NM613304, altitude 200 m OD) and Torness (grid ref. NM643333, altitude 110 m OD) are located in Glen More within the mapped limits of Loch Lomond Stadial ice, and these infilled kettle-hole basins contain only Flandrian limnic and terrestrial sediments. At Loch an t-Suidhe and Mishnish, samples for pollen analysis and radiocarbon dating were removed from a deep part of the basins using a piston corer of 5 cm diameter. At the sites in Glen More, however, it was necessary to ensure that the earliest sediments that had accumulated following the wastage of Loch Lomond Stadial ice were being sampled, and thus the infills along a 5 m by 10 m grid pattern were sounded to establish the subsurface contours of the basins⁹. Cores were then taken from the deepest point. At all the sites, samples for radiocarbon dating were obtained by bulking gyttja material from comparable biostratigraphical horizons in up to six closely-spaced cores, the precise levels from which individual samples were to be taken being determined on the basis of relative pollen counts in each core. In this way, narrow bands of sediment 2–5 cm in thickness could be dated. The radiocarbon dates obtained and the principal characteristics of the dated horizons are presented in Table 1.

Difficulties in the dating of Lateglacial and early Flandrian pollen-stratigraphical records in limnic sediments arise because of the possibility of (1) 'hard water error'¹⁰ particularly in dated samples from newly-deglaciated terrain¹¹; (2) contamination by younger carbon where the sediments have been affected by percolating groundwaters^{12,13}; (3) delayed ice-melt in kettle-holes which can produce a basal biostratigraphical record and corresponding radiocarbon dates that postdate regional deglaciation at some sites by hundreds or even thousands of years^{14,15}; (4) relatively slow rates of sediment accumulation in many basins at times of rapid environmental change which result in poor levels of stratigraphical resolution for the dating of key pollen-assemblage zones¹⁶. There are also uncertainties associated with the radiocarbon method. As many of the potential error sources cannot be quantified at individual sites, it is suggested that the dating of Lateglacial and early Flandrian events within any one area should be based on radiocarbon dates from different biostratigraphical horizons. An indication of the 'true' radiocarbon age of each horizon can then be obtained by taking the mean of the individual values. This procedure is followed where possible for the new radiocarbon dates reported here. In averaging the dates from Mull it is, of course, assumed that no significant time-transgression is associated with the vegetational developments identified at each site, a reasonable assumption in view of the relatively small size of the island.

Figure 2 shows the dates listed in Table 1 plotted with one standard deviation about the mean in relation to the principal biostratigraphical and lithostratigraphical horizons. The evidence suggests the following chronostratigraphical framework for environmental changes:

(1) *Corylus* rise: two dates obtained for this event show remarkably close agreement, averaging ~8,800 yr BP. Radiocarbon dates on the *Corylus* expansion at sites on the Isle of Skye range from 9,655 to 7,500 yr BP, although the latter age determination was considered to be aberrant^{17,18}. However, comparable dates to these from Mull have been obtained from sites in Wester Ross (8,951 ± 120 yr BP¹⁹) and the Spey Valley (8,670 ± 150 yr BP)²⁰. The available evidence points to the widespread colonization by hazel of the Scottish Highlands and the islands of the Inner Hebrides shortly after 9,000 yr BP. The dates from Skye might indicate local exceptions to this generalization, but this requires further examination. The expansion of hazel on Mull appears to postdate that in lowland Scotland by some 300–400 yr (ref. 21).

(2) *Juniperus* maximum and *Betula* expansion: the two dates obtained for the early Flandrian *Juniperus* maximum average

~9,600 yr BP, while the single date obtained for the *Betula* expansion ($9,350 \pm 70$ yr BP) is in good stratigraphical agreement with the former and also with the dates for the later *Corylus* rise.

(3) Loch Lomond Stadial–early Flandrian transition: biostratigraphical resolution of the Loch Lomond Stadial–early Flandrian transition in Scottish sites is seldom clear because of the often slow rates of sediment accumulation, and the rapid vegetational changes that occurred between climatic amelioration at the close of the Loch Lomond Stadial and the expansion of juniper scrub. In addition, it is often impossible to develop a chronology based on radiocarbon which differentiates between the individual pollen–stratigraphical phases within this time period. The dates from Mull on all pre-*Juniperus* horizons above the Loch Lomond Stadial–early Flandrian lithostratigraphical boundary have, therefore, been grouped together. Of the six dates shown in Fig. 2, that from Coire Clachach ($9,530 \pm 80$ BP) is clearly aberrant. The others range between $10,000 \pm 70$ yr BP and $10,440 \pm 80$ yr BP suggesting a mean age of 10,200 yr BP for the establishment of more stable conditions on Mull following the harsh environment of the Loch Lomond Stadial. The dates from Loch an t-Suidhe are considered to be of particular significance. Sedimentation in this basin seems to have been unusually rapid during the Loch Lomond Stadial–early Flandrian transition, and it was possible to obtain three radiocarbon dates (SRR 1800–1802) from biostratigraphical horizons preceding the *Juniperus* phase. The fact that these dates are closely grouped, internally consistent and in broad agreement with a number of dates from basal Flandrian sediments from other Scottish sites^{22,23} reinforces the suggestion^{24,25} that the transition to more stable environmental conditions in Scotland at the end of the Loch Lomond Stadial occurred well before 10,000 yr BP.

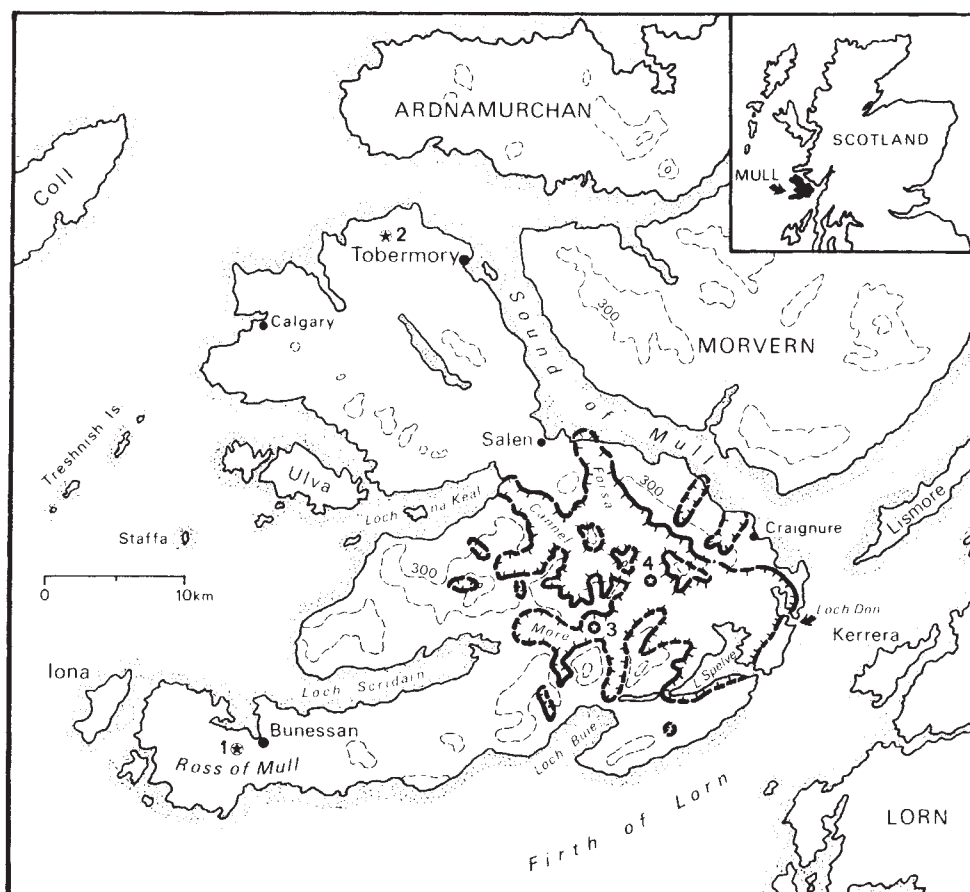
(4) Close of the Lateglacial Interstadial: this event is dated on the Isle of Mull to ~10,700 yr BP on the basis of two age determinations which are in remarkably close agreement. The date is younger than the age assigned to the lower boundary

Table 1 Radiocarbon dates from the Isle of Mull

Site	Sample no.	Depth (m)	Thickness (cm)	Date (yr BP)
Loch an t-Suidhe	SRR-1800	7.74	5	$10,200 \pm 70$
Loch an t-Suidhe	SRR-1801	7.80	5	$10,200 \pm 70$
Loch an t-Suidhe	SRR-1802	7.85	5	$10,440 \pm 80$
Loch an t-Suidhe	SRR-1803	8.36	5	$10,690 \pm 70$
Loch an t-Suidhe	SRR-1804	8.63	5	$11,860 \pm 80$
Loch an t-Suidhe	SRR-1805	8.77	5	$13,140 \pm 100$
Mishnish	SRR-1806	4.96	4	$10,000 \pm 70$
Mishnish	SRR-1807	5.08	4	$10,730 \pm 60$
Coire Clachach	SRR-1594	5.80	5	$8,870 \pm 120$
Coire Clachach	SRR-1595	6.00	5	$9,350 \pm 70$
Coire Clachach	SRR-1596	6.65	5	$9,500 \pm 70$
Coire Clachach	SRR-1597	6.70	5	$9,530 \pm 80$
Torness	SRR-1799	6.65	2	$8,760 \pm 140$
Torness	SRR-1798	6.67	2	$9,660 \pm 140$
Torness	SRR-1797	6.69	2	$10,170 \pm 150$

of the Younger Dryas chronozone in north-west Europe⁶, and is also younger than age determinations on comparable horizons at several sites on the Scottish mainland^{18,20,26}. One possibility is that this reflects the more westerly location of Mull and its exposure to moderating maritime influences which could have produced a delayed geomorphological and vegetational response to climatic deterioration, the effects of which were reflected earlier in the pollen stratigraphy at sites on the Scottish mainland. However, it is now generally accepted that climatic deterioration at the onset of the Loch Lomond Stadial was a result of oceanographic changes in the North Atlantic^{27,28}, in particular the southward migration of the oceanic (and atmospheric) polar front, and therefore climatic deterioration might be expected earlier in the west and not later as is implied by the radiocarbon dates. Yet, at other sites in eastern and central Scotland, dates have been obtained from Lateglacial profiles which are comparable with those reported here from

Fig. 1 The Isle of Mull. Pollen sites: 1, Loch an t-Suidhe; 2, Mishnish; 3, Coire Clachach; 4, Torness. The heavy lines show the limits of Loch Lomond Stadial ice (after ref. 5). Contour interval 300 m.



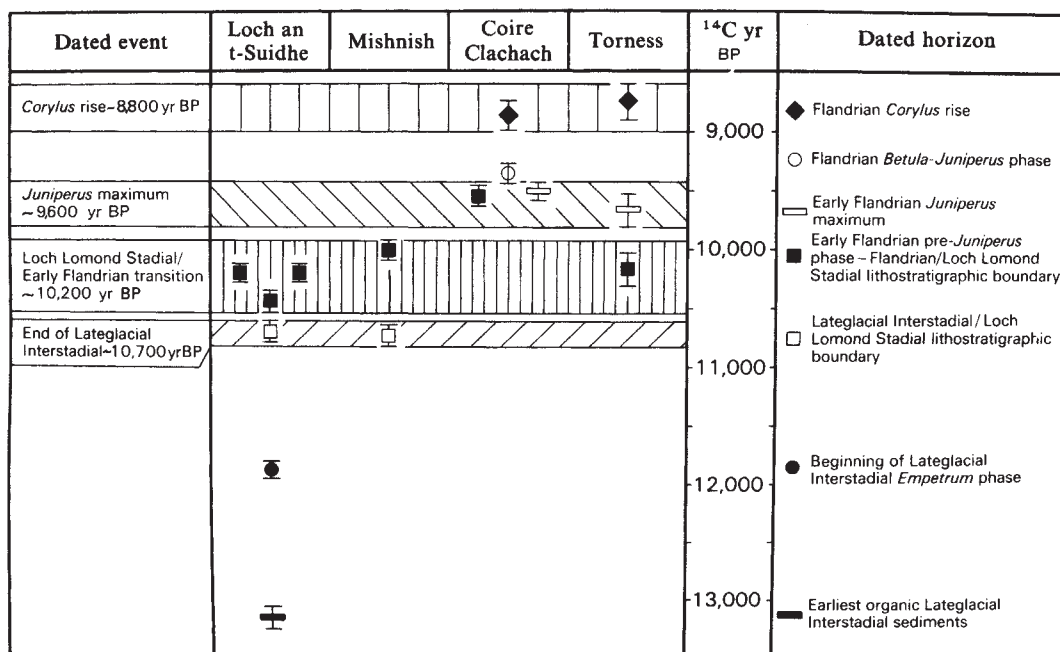


Fig. 2 Radiocarbon dating of Lateglacial and early Flandrian environmental changes at the four sites on the Isle of Mull.

Mull^{18,29,30}. Clearly, therefore, there is no consistent spatial pattern in Scottish radiocarbon dates for the close of the Late glacial Interstadial, and it may be that local site factors rather than the macroscale climatic changes were of greater importance in determining the precise nature of lake basin biostratigraphy at the Lateglacial Interstadial-Loch Lomond Stadial transition.

(5) Lateglacial Interstadial: two dates were obtained from the Lateglacial Interstadial deposits at Loch an t-Suidhe. The immigration of *Empetrum* and, to a lesser extent, *Juniperus* and *Betula*, was dated at $11,860 \pm 80$ yr BP, while the basal organic sediments yielded a radiocarbon age of $13,140 \pm 100$ yr BP. A comparison between the three dates from the Lateglacial Interstadial at Loch an t-Suidhe (SSR 1803-1805) in relation to their position in the stratigraphical column (Table 1) indicates that the rate of sediment accumulation in the basin during the later part of the Interstadial was almost double what it had been in the earlier stages. Although sedimentation in lakes can be highly variable as a result, for instance, of different rates of inwash and of organic productivity, such a marked discrepancy is unusual, and may reflect errors in one or more of the radiocarbon dates. The age determination of $11,860 \pm 80$ yr BP is difficult to evaluate in the context of the Scottish Lateglacial because of the marked contrasts which appear to have existed in vegetational developments between Mull and the mainland, particularly in respect of the reduced representation of woody taxa in the former area. Moreover, very few dates have so far been published on vegetational changes during the Lateglacial Interstadial in Scotland. The pollen spectra associated with the basal date (SRR 1805) are characterized by a *Rumex-Salix*-Gramineae-Cyperaceae assemblage and reflect an early stage in plant colonization following the disappearance of the Devensian ice sheet. Although a hardwater error cannot be excluded¹¹, the date of $13,140 \pm 100$ yr BP is very similar to age determinations on comparable biostratigraphical horizons at sites on the Scottish mainland^{18,30-33}, and is also in agreement with the inferred date for the replacement of polar by warmer waters around the coasts of western Britain²⁷. If correct, therefore, it would support the suggestion that climatic amelioration at the beginning of the Lateglacial Interstadial in Scotland began around, or even before, 13,000 yr BP (ref. 34).

We conclude with the following framework for environmental change. Wastage of the Scottish ice sheet was followed by climatic amelioration, perhaps around 13,000 yr BP, after which Mull gradually developed a vegetation of *Empetrum* heath, juniper and birch scrub and open grassland. Climatic deterior-

ation at, or before 10,700 yr BP led to the breakup of the Interstadial vegetation cover, the development of a local ice cap and several separate smaller glaciers in the hills of central Mull, and the establishment of severe periglacial conditions throughout the island. Climatic amelioration around 10,200 yr BP caused the final wastage of the Loch Lomond Stadial glaciers and initiated the early Flandrian plant succession. This began with an *Empetrum* phase and was succeeded by *Juniperus* expansion between 9,600 and 9,500 yr BP, the establishment of open birchwoods by ~9,300 yr BP and the immigration of *Corylus* around 8,800 yr BP. The above chronology is strengthened by the internal consistency of the Mull dates and by the broad measure of agreement between dates from similar biostratigraphical horizons, although clearly the precise timing of events during the early Lateglacial Interstadial remains enigmatic.

This research was financed by the NERC, by the Pantyfedwen Fund of St David's University College, Lampeter, and by the City of London Polytechnic. We thank Dr D. D. Harkness for the radiocarbon dates, Dr R. Cornish, Dr A. G. Dawson, Dr J. M. Gray and Mr. D. John for assistance with the fieldwork.

Received 5 November 1981; accepted 5 February 1982.

1. Bailey, G. B., Clough, C. T., Wright, W. B., Richey, J. E. & Wilson, G. V. *Mem. geol. Surv. U.K.* (1924).
2. Ballantyne, C. K. & Wain-Hobson, T. *Scott. J. Geol.* **16**, 1-10 (1980).
3. Sissons, J. B. *Scott. J. Geol.* **13**, 23-36 (1977).
4. Sissons, J. B. *The Evolution of Scotland's Scenery* (Oliver & Boyd, Edinburgh, 1967).
5. Gray, J. M. & Brooks, C. L. *Scott. J. Geol.* **8**, 95-103 (1972).
6. Mangerud, J., Anderson, S. T., Berglund, B. E. & Donner, J. J. *Boreas* **3**, 109-126 (1974).
7. Lowe, J. J. & Gray, J. M. in *Studies in the Lateglacial of North-West Europe* (eds Lowe, J. J., Gray, J. M. & Robinson, J. E.) 157-175 (Pergamon, Oxford, 1980).
8. Sissons, J. B., Lowe, J. J., Thompson, K. S. R. & Walker, M. J. C. *Nature phys. Sci.* **244**, 75-77 (1973).
9. Walker, M. J. C. & Lowe, J. J. *J. Biogeogr.* **4**, 333-351 (1977).
10. Shotton, F. W. *Nature* **240**, 460-461 (1972).
11. Sutherland, D. G. in *Studies in the Lateglacial of North-West Europe* (eds Lowe, J. J., Gray, J. M. & Robinson, J. E.) 139-149 (Pergamon, Oxford, 1980).
12. Lowe, J. J. & Walker, M. J. C. in *Studies in the Lateglacial of North-West Europe* (eds Lowe, J. J., Gray, J. M. & Robinson, J. E.) 123-137 (Pergamon, Oxford, 1980).
13. Lowe, J. J. & Walker, M. J. C. *Boreas* **10**, 281-294 (1981).
14. Porter, S. C. & Carson, R. J. *Quat. Res.* **1**, 410-414 (1971).
15. Post, A. *Quat. Res.* **6**, 111-118 (1976).
16. Lowe, S. *Nature* **293**, 210-212 (1981).
17. Birks, H. J. B. *The Past and Present Vegetation of the Isle of Skye: a Palaeoecological Study* (Cambridge University Press, 1973).
18. Vasari, Y. in *Studies in the Scottish Lateglacial Environment* (eds Gray, J. M. & Lowe, J. J.) 143-162 (Pergamon, Oxford, 1977).
19. Birks, H. H. *New Phytol.* **71**, 731-754 (1972).
20. Birks, H. H. & Mathewes, R. W. *New Phytol.* **80**, 455-484 (1978).
21. Lowe, J. J. *Pollen Spores* **20**, 367-397 (1978).
22. Walker, M. J. C. & Lowe, J. J. in *Timescales in Geomorphology* (eds Cullingford, R. A., Davidson, D. A. & Lewin, J.) 247-259 (Wiley, New York, 1980).
23. Walker, M. J. C. & Lowe, J. J. *J. Biogeogr.* **6**, 349-361 (1979).

24. Lowe, J. J. & Walker, M. J. C. *Nature* **264**, 632–633 (1976).
25. Walker, M. J. C. & Lowe, J. J. *J. Biogeogr.* **8**, 475–491 (1981).
26. Lowe, J. J. & Walker, M. J. C. in *Studies in the Scottish Lateglacial Environment* (eds Gray, J. M. & Lowe, J. J.) 101–118 (Pergamon, Oxford, 1977).
27. Ruddiman, W. F. & McIntyre, A. *Quat. Res.* **3**, 117–130 (1973).
28. Ruddiman, W. F., Sancetta, C. D. & McIntyre, A. *Phil. Trans. R. Soc. B* **280**, 119–142 (1977).
29. Godwin, H. & Willis, E. H. *Proc. R. Soc. B* **150**, 199–215 (1959).
30. Sissons, J. B. & Walker, M. J. C. *Nature* **249**, 822–824 (1974).
31. Kirk, W. & Godwin, H. *Trans. R. Soc. Edinb.* **65**, 225–249 (1963).
32. Bishop, W. W. *Trans. J. Proc. Dumfries. Galloway nat. Hist. Antiq. Soc.* **40**, 117–132 (1963).
33. Pennington, W. *Boreas* **4**, 157–171 (1975).
34. Bishop, W. W. & Coope, G. R. in *Studies in the Scottish Lateglacial Environment* (eds Gray, J. M. & Lowe, J. J.) 61–88 (Pergamon, Oxford, 1977).

Life span of the biosphere

J. E. Lovelock

Coombe Mill, St Giles on the Heath, Launceston,
Cornwall PL15 9EY, UK

M. Whitfield

Marine Biological Association of the United Kingdom,
The Laboratory, Citadel Hill, Plymouth PL1 2PB, UK

There has been life on Earth for at least 3,500 Myr but the assumption that a comparable future lies ahead may not be justified. Main sequence stars appear to increase their burning rate as they age. Thus the Sun, if a typical star, can be predicted to have increased its output by 30% since the Earth's origin 4,500 Myr ago¹. The maintainance of an equable climate since life began probably required some means of planetary thermostasis. The Gaia hypothesis proposed by Lovelock and Margulis² included an unspecified biological means for climate control. Walker *et al.*³ suggests an abiological automatic thermostasis in which the atmospheric abundance of CO₂, a greenhouse gas, adjusts to resist the warming tendency of the increased solar flux. Here we discuss possible links between the biological and geological control mechanisms. It is clear that whatever the mechanism, atmospheric CO₂ is now close to its lower limit of partial pressure, so the biosphere may soon, in geological terms, be exposed without protection to the predicted progressive increase of solar luminosity.

In spite of the probability of a 25% increase in solar luminosity it is reasonably certain that the Earth's climate has undergone little change since life began. Both the geological record and the persistence of life indicate that neither global freezing nor boiling conditions have ever prevailed. Indeed, the Earth's mean surface temperature has probably never departed from the range 5–50 °C and, contrary to the predictions of simple physics, may have been warmest at the beginning when the Sun was small and has cooled ever since⁴.

Several models have been proposed to account for the Earth's relatively constant temperature. Sagan and Mullen⁵ were the first to suggest that the infant biosphere was warmed by an atmospheric gas which exerted a 'greenhouse' effect by transmitting sunlight while hindering the escape of heat to space.

Although water vapour makes the most significant contribution to the greenhouse effect in the contemporary atmosphere, because of its physical properties water is an unlikely candidate for long-term thermostasis. Its relatively high freezing and boiling points render its blanketing effect prone to positive feedback in the face of extreme temperature excursions. For example, a sudden fall in temperature could result in an increase in the size of the polar ice caps and the seasonal snow fields (thus increasing the efficiency with which radiation is reflected from the Earth) and a corresponding fall in the atmospheric humidity. Both effects could contribute to a further drop in temperature. On the other hand, a sudden rise in temperature would increase the water vapour content of the atmosphere, which would in turn push the temperature still higher. In consequence, the complexities of cloud cover feedback effects over

long periods of time have focused attention on simpler atmospheric models involving other greenhouse gases^{3–5}.

Sagan and Mullen⁵ suggested that the early atmosphere was rich in ammonia and other reduced gases and that these provided the blanket which kept the Earth sufficiently warm for life to emerge. Relatively high partial pressures of ammonia were also seen as essential for the chemical evolution of life and, as ammonia is rapidly photolysed in the atmosphere, a significant inorganic source was sought⁶. Recent work, however, suggests that high partial pressures of ammonia are not essential for the abiological production of amino acids⁷ and that the primordial atmosphere probably contained little ammonia but relatively high partial pressures of CO₂⁸. The concept of an early Earth warmed by the blanketing effect of its atmosphere is still therefore feasible but with CO₂ the preferred blanket gas^{4,7}. According to this hypothesis whatever greenhouse gas or other agency kept the young planet warm, it must have been smoothly and actively reduced from then until now; otherwise the mean temperature would have progressively increased with increasing solar heat flux and might by now have exceeded 50 °C, the critical upper limit for most life⁹.

Walker *et al.*³ have proposed that the climate could have been controlled solely by abiological negative feedback involving a gradual decline in the atmospheric partial pressure of CO₂ in response to a continuous increase of the Sun's luminosity. Briefly, an increase in temperature accelerates the rate of reaction between CO₂ and calcium silicate rock. The input of CO₂ from tectonic sources is assumed to be constant so that an increased heat flux leads to increased weathering of the rock, a faster removal of CO₂ and hence a lower atmospheric CO₂ partial pressure. This cybernetic process acts to resist the rise in temperature which would otherwise result from the increase in solar luminosity.

This abiological mechanism could have operated early in the Earth's history but the resulting climatic control would have been easily perturbed by fluctuations in the volcanogenic flux of CO₂. These effects could have been exaggerated by the positive feedback loops introduced by the ocean-atmosphere interaction. As the solubility of CO₂ in water decreases with increasing temperature, an atmospheric (and hence oceanic) warming would result in a release of CO₂ to the atmosphere which would produce a further temperature rise. A corresponding positive feedback would occur on cooling. At present geological input and removal constitute only a few tenths of a per cent of the biologically driven fluxes¹⁰ and so it is instructive to see how the biological and geological mechanisms can be linked.

There is little doubt that current weathering of silicate rocks is biologically, not geochemically, determined. The partial

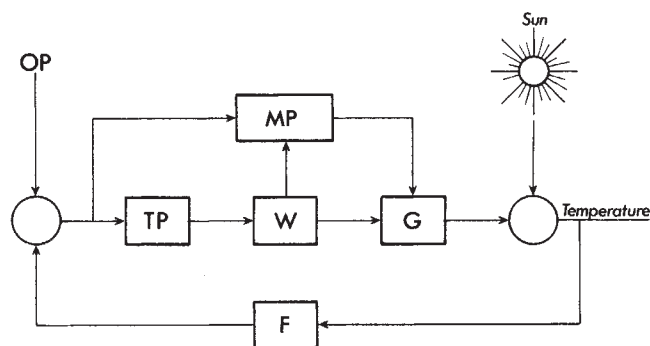


Fig. 1 A feedback process in which the productivity of the terrestrial biosphere (TP) or the marine biosphere (MP) influences the greenhouse effect (G) and hence the global mean temperature via control of the atmospheric partial pressure of carbon dioxide. These biological processes modulate the geological control exerted by the weathering process (W) against a constant input from tectonic sources. The loop is closed by the feedback process (F, see text) which transduces from the mean temperature a productivity that can be compared with the operating productivity (OP) of the biosphere.

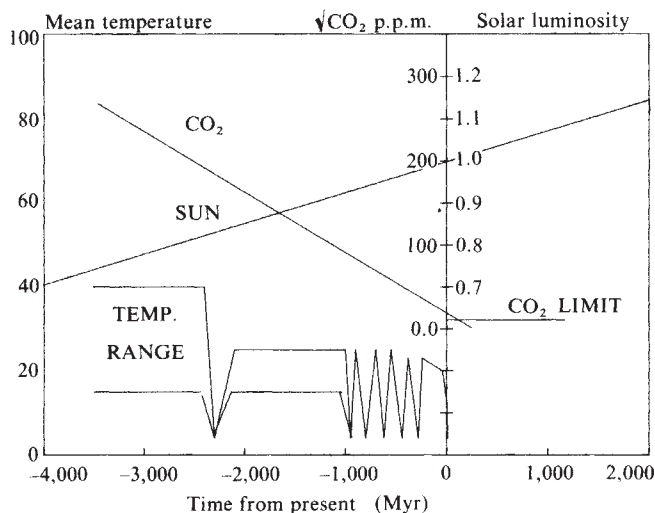


Fig. 2 Evolution of the climate showing the variation of solar luminosity, relative to its present value taken as 1.0. Also illustrated on the same time scale are the proposed decline in carbon dioxide concentration, expressed as the square root of the concentration in parts per million (p.p.m.) by volume and the approximate range of mean surface temperatures in °C (see text).

pressure of CO_2 in the soil where the chemical weathering takes place is 10–40 times higher than the atmospheric level¹¹. These high partial pressures are maintained by the biological oxidation of organic detritus and the rate of oxidation approximately doubles for each 10 °C rise in temperature. The biota seem to act as a sensor of temperature change and as an amplifier to magnify the rate of weathering of the silicate rocks and pumping of CO_2 from the air. The present geochemical balance for CO_2 between input from tectonic sources and removal by weathering¹⁰ requires a partial pressure in the soil of nearly 10 mbar.

The direct involvement of the biota in weathering would only be possible after the evolution of a land microfauna and flora and the establishment of viable soil communities. Although such communities probably predate the invasion of land as deduced from fossils, it is possible that life was largely confined to the sea during the Archaean. Once communities based on photosynthetic carbon fixation were established in the oceans, however, they could amplify indirectly the geochemical feedback. In the contemporary ocean the biota limit the atmospheric partial pressure of CO_2 by maintaining a continuous flux of particulate organic matter into the deep ocean¹². This flux results from primary production in the surface layers which is limited by the rate of supply of nutrients, notably nitrogen and phosphorus, from riverine inputs and from the slow recirculation of nutrient- and carbon dioxide-rich deep-ocean water. If, as suggested by the geological model³, the rate of rock weathering increases due to a rise in global temperature, then the rate of supply of nutrient elements to the oceans will also increase. This will enhance primary productivity and further reduce atmospheric CO_2 levels. If, on the other hand, the temperature falls the nutrient supply will diminish, productivity will decrease and excess CO_2 will be released from the recirculating deep-ocean water into the atmosphere, thus encouraging a temperature rise¹². Here, too, the biota can use geologically driven processes to provide additional temperature stabilization. If life ceased now, the new geochemical balance would require an increase of atmospheric CO_2 and hence temperature until removal by weathering again equalled the input of CO_2 from volcanoes. The present low atmospheric CO_2 pressure and relatively low temperature are therefore a direct consequence of life. This is consistent with the Gaia hypothesis² which suggests that evolutionary pressures ensure that biologically driven processes which influence the chemical composition of the atmosphere and hence the climate so as to keep the environ-

ment optimal for the biosphere will have been preferentially selected.

Figure 1 illustrates a highly simplified model of the combined biological and geological regulation of temperature. As the temperature falls, CO_2 fixation by photosynthesis and chemical weathering declines until at ~0 °C it nearly ceases. CO_2 input from tectonic sources continues and the atmospheric partial pressure of CO_2 will rise. Through the greenhouse effect, this increase resists any further decline in temperature. In a similar way high temperatures and high CO_2 pressures are both conducive to rapid weathering which serves to resist a continuing rise in temperature. Thus the mean surface temperature will tend to a value optimal for the contemporary biota within the other constraints of its planetary environment.

Figure 2 illustrates the probable solar output from 4,000 Myr ago to 2,000 Myr hence and gives an estimate of the corresponding global mean temperature and CO_2 partial pressure. Solar luminosity is predicted to increase, according to Newkirk¹. An alternative analytical expression for the time variation of solar luminosity¹³ might be useful for more detailed calculations.

The range of past temperatures is not accurately known. That shown in Fig. 2 is our estimate of the probable limits of palaeoclimates and it is based on the following assumptions. (1) During the Archaean the biosphere was wholly prokaryotic. As the upper temperature limit for contemporary prokaryotes is 50 °C, the upper limit for the Earth's mean temperature was set lower, at 40 °C, to allow for zonal and other variations. No glaciations are known during this period^{14,15} so the lower limit of temperature is set at 15 °C, close to the present value. (2) –2,200 Myr marks the end of the Archaean and the first extensive glaciation^{14,15}, as well as the appearance of free oxygen. Temperatures might have risen after this glaciation to Archaean levels or they may have been similar to the levels now thought characteristic of the interglacial periods—20–30 °C. We chose the latter estimate as the upper limit from –2.2 Gyr until the present. The lower limit is set by the physical properties of water. There is no indication that total freezing took place, so a mean temperature of 0 °C or below is very unlikely. Glaciations^{14,15} are associated with mean temperatures of 5–10 °C and interglacials with lower limits perhaps 5–15 °C higher. The time scale of the figure is so compressed that the present interglacial climate cannot be resolved from the low values of the last glaciation 10,000 yr ago.

The other quantity illustrated in Fig. 2 is the CO_2 pressure needed to compensate for the changing solar luminosity. This estimate is based on a partial pressure 3,500 Myr ago of 7,000 p.p.m. (refs 4, 17). It is linked to the present value of 320 p.p.m. on the assumption of a linear relationship between the square root of the CO_2 pressure and the solar luminosity for a constant mean temperature. The exact relationship may be different, however, and an analysis of the geological mechanism³ suggests that a 2/3 power relationship might be more appropriate. Nonetheless, such differences are unlikely to affect the qualitative conclusions that can be drawn from Fig. 2.

If the principal agent for temperature regulation is the partial pressure of CO_2 in the atmosphere and even if the solar luminosity is constant, it seems that we are very close to the lower limit of possible adjustment; also there are no other gases in the air which could serve significantly in this way by the reduction of their partial pressures. Other mechanisms such as an increase in albedo by desert cover might sustain an equable climate but as observed by Henderson-Sellers¹⁶, this is generally neutralized by a concomitant change in cloud cover. On the other hand, if we accept the predicted increase in luminosity and assume that the future decline of CO_2 matches the solar output as may have been the case in the past, then 150 p.p.m. pressure will be reached in about 100 Myr. This concentration is the lower limit tolerable for photosynthesis. Some adaptation to lower CO_2 concentrations and to higher temperatures is possible but it would not buy much time. In human terms the crisis is still infinitely distant but in terms of the life span of the biosphere,

rich with familiar metazoans, we might forecast an end to the long spell of cool and favourable climate.

We thank Dr A. F. Tuck and Professor R. M. Garrels for stimulating suggestions.

Received 4 June 1981; accepted 12 February 1982.

1. Newkirk, G. Jr. *Geochim. cosmochim. Acta Suppl.* **13**, 293 (1980).
2. Lovelock, J. E. & Margulis, L. *Tellus* **26**, 1 (1974).
3. Walker, J. C. G., Hays, P. B. & Kasting, J. F. *J. geophys. Res.* (in the press).
4. Owen, T., Cess, R. D. & Ramanathan, V. *Nature* **277**, 640 (1979).
5. Sagan, C. & Mullen, G. *Science* **177**, 52 (1972).
6. Henderson-Sellers, A. & Schwarz, A. *Nature* **287**, 526 (1980).
7. Wigley, T. M. L. & Brimblecombe, P. *Nature* **291**, 213 (1981).
8. Henderson-Sellers, A., Benlow, A. & Meadows, A. J. *Q. Jl R. astr. Soc.* **21**, 74 (1980).
9. Brock, T. D. *Science* **158**, 1012 (1967).
10. Bolin, B., Degens, E. T., Kempe, S. & Ketner, P. *The Global Carbon Cycle*, SCOPE Rep. 13 (Wiley, New York, 1979).
11. Holland, H. D. *The Chemistry of the Atmosphere and Oceans* (Wiley, Chichester, 1978).
12. Baes, C. F. in *Carbon Dioxide Review 1982* (ed. Clark, W. C.) (Institute of Energy Analysis, Oak Ridge Associated Universities, Tennessee, 1982).
13. Gough, D. in *The Solar Output and its Variation* (ed. White, O. R.) 451 (Colorado Associated University Press, 1977).
14. Steiner, J. & Grillmair, E. *Bull. geol. Soc. Am.* **84**, 1003 (1973).
15. Harland, W. B. & Herod, K. N. in *Ice Ages Ancient and Modern* (eds Wright, A. E. & Moseley, F.) (Seal House, Liverpool, 1975).
16. Henderson-Sellers, A. *Nature* **279**, 786 (1979).
17. Hart, M. H. *Icarus* **33**, 23 (1978).

LDH-B genotype-specific hatching times of *Fundulus heteroclitus* embryos

Leonard DiMichele & Dennis A. Powers

Department of Biology, The Johns Hopkins University, Baltimore, Maryland 21218, USA

The evolutionary significance of protein polymorphisms has long been debated. Exponents of the balanced theory advocate that selection operates to maintain polymorphisms, while those of the neoclassical school argue that most genetic variability is selectively neutral¹. As the 'neutralist' hypothesis implies that allelic isozymes are functionally equivalent, some investigators have examined the biochemistry of protein polymorphisms²⁻¹⁵, while others have concentrated on life history correlates¹⁶⁻²³. Few studies, however, have established that *in vitro* functional differences are reflected at the whole organism level^{5,21-23}, yet this is a critical link in understanding the significance of protein polymorphisms. We have studied the effects of the kinetically different lactate dehydrogenase-B (LDH-B) allelic isozymes on the rate of development and physiological performance of the fish, *Fundulus heteroclitus*, and report here that hatching time is highly correlated with LDH-B genotype: *LDH-B^aB^a* individuals hatch before *LDH-B^bB^b* fish while heterozygote (*LDH-B^aB^b*) hatching is intermediate. The basis for this phenomenon may be a differential ability to deliver oxygen to respiring tissues. As precisely timed hatching is critical to the survival of *Fundulus*²⁴, such differences in hatching time between LDH-B genotypes may be an important component of the species' evolutionary strategy.

The LDH-B polymorphism in *F. heteroclitus* is due to two co-dominant alleles^{25,26} which exhibit a cline in gene frequencies along the Atlantic Coast of the United States^{4,27}. We have shown that there are temperature- and pH-dependent kinetic differences between the LDH-B allelic isozymes⁷. Moreover, intra-erythrocyte ATP differences among individuals are correlated with LDH-B phenotype³. The role of ATP as an allosteric modifier of *Fundulus* haemoglobin is similar to that of 2,3-diphosphoglycerate in human red blood cells, in that it decreases the affinity of haemoglobin for oxygen^{28,29}. Thus, fish with the highest levels of intra-erythrocyte ATP (that is, *LDH-B^bB^b*) have the lowest blood oxygen affinity³. As respiratory stress has been shown^{30,31} to trigger the hatching mechanism in *Fundulus*, we hypothesized that the hatching times of LDH-B genotypes should differ because of differences in blood oxygen affinity. Specifically, *LDH-B^aB^a* fish should hatch before *LDH-B^bB^b* individuals.

As an initial test of this hypothesis, fertilized eggs were obtained by stripping gametes from males and females that were randomly paired. Each of the resulting clutches was incubated in a separate Petri dish at 20°C, pH 7.0 and 15‰ salinity (Instant Ocean). The incubation medium was changed daily and the fry were collected as they hatched. The LDH-B genotypes of each parental pair and their offspring were determined by gel electrophoresis^{26,27} (see Fig. 1). Clearly, hatching of *LDH-B^aB^a* eggs was predominant in the first 3 days, while *LDH-B^bB^b* eggs predominated in the last 3 days. The overall mean hatching times were 11.9, 12.4 and 12.8 days for the *LDH-B^aB^a*, *LDH-B^aB^b* and *LDH-B^bB^b* genotypes, respectively. Individual crosses in which there were more than one genotype also showed that *LDH-B^aB^a* eggs hatched before *LDH-B^bB^b*, which in turn hatched before *LDH-B^aB^b* (Fig. 1).

In addition to establishing an ordered hatching of LDH-B genotypes, our analysis (Fig. 1) also indicated variations between crosses that were not due to the LDH-B locus. If these variations were due to other loci, then pooling gametes from many individuals before fertilization should dramatically reduce the variation between crosses. To test this hypothesis, we performed experiments in which adult fish were segregated by LDH-B genotype into breeding stocks²⁶. Specific LDH-B crosses were made by pooling gametes before fertilization. Consistent with our expectation, these experiments (Tables 1 and 2) almost eliminated variation between similar crosses. Thus, some of the variation among the 20 random crosses in Fig. 1 must have been due to contributions from other loci. On the other hand, Tables 1 and 2 indicate that there are still highly significant differences ($P < 0.01$) in hatching time between LDH-B genotypes. These data (Tables 1 and 2), together with

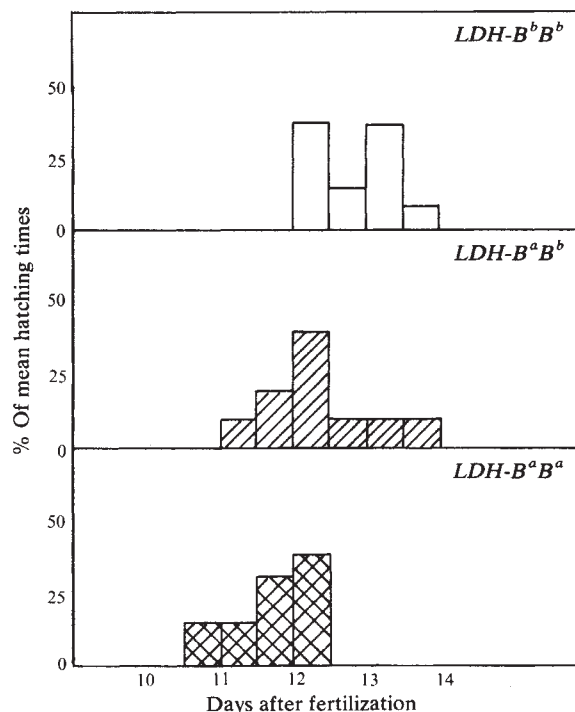


Fig. 1 Distribution of mean hatching times for three LDH-B genotypes from 20 random crosses. Analysis of variance³⁶ indicated significant differences between the crosses ($P < 0.05$); Duncan's multiple range test³⁶ showed that some of this variation was due to LDH-B genotype (for example, all the *LDH-B^aB^a* × *LDH-B^aB^a* crosses hatched before all the *LDH-B^bB^b* × *LDH-B^bB^b* crosses; $P < 0.05$). There were also significant differences, however, between similar crosses (for example, two of the *LDH-B^aB^a* × *LDH-B^aB^a* crosses hatched sooner than three other crosses of the same type; $P < 0.05$). The histogram shows the distribution of mean hatching times among genotypes irrespective of parental cross. Separate analyses of each cross in which there were offspring of more than one genotype, showed significant differences ($P < 0.05$) in hatching time between genotypes in each case.

Table 1 Mean hatching times (days) for each offspring genotype resulting from four $LDH-B^aB^b \times LDH-B^aB^b$ 'mass' crosses

Replicate	N	Offspring genotype*		
		$LDH-B^aB^a$	$LDH-B^aB^b$	$LDH-B^bB^b$
1	2,142	12.0	12.8	12.9
2	1,273	11.9	12.0	13.1
3	1,396	12.4	12.8	14.8
4	1,051	12.5	12.5	14.3

Sperm and eggs from 40 $LDH-B^aB^b$ males and 40 $LDH-B^aB^b$ females were pooled in separate dishes and then mixed to make each cross. N, total number of offspring from each cross.

* The data were analysed by a two-way analysis of variance of unweighted means³⁶, which provides a conservative estimate of the variance between genotypes. The differences between genotypes was found to be significant ($P < 0.01$); there were also significant differences between replicates ($P < 0.05$) but this estimate is not conservative because the MS_{error} term had to be calculated separately.

the consistent trend from the 20 random crosses (Fig. 1), demonstrate that hatching in *F. heteroclitus* occurs generally in the following order: $LDH-B^aB^a$ phenotype before $LDH-B^aB^b$, then $LDH-B^bB^b$.

The eggs of *F. heteroclitus* are laid in empty mussel shells³², between the leaves of the marsh grass, *Spartina alterniflora*²⁴, or in the sand above the mean high tide³⁵. In these conditions, the eggs incubate in air for most of their developmental period. Hatching occurs when eggs laid at one spring tide are immersed in water by the following spring tide. As water covers the eggs, there is a drop in environmental oxygen at the egg surface which is the hatching cue for the embryo³⁰. Therefore, overall plasticity in hatching times may be important in protecting *F. heteroclitus* populations that live in variable environmental conditions. Our data suggest that premature hatching cues (for example, rainstorms) would select largely against $LDH-B^aB^a$ individuals, while late hatching (that is, after the tide has retreated) would select primarily against the $LDH-B^bB^b$ phenotype. This argument is particularly compelling in view of the finding of Meredith and Lotrich³⁴ that the mortality of *F. heteroclitus* in age-class 0 (eggs to fry of 59 mm) is $>99.5\%$.

As respiratory stress initiates hatching and blood oxygen affinity is correlated with $LDH-B$ genotype via intra-erythrocyte ATP concentrations, the simplest interpretation of our data is that the differences in hatching times result from functional differences between $LDH-B$ allelic isozymes. However, our results do not exclude the possibility that other genes, tightly linked to the $LDH-B$ locus, may be responsible for the observed effects on hatching time. Whatever the molecular basis for these differences may be, it is clear that genetic variability at the $LDH-B$ locus (or a closely linked one) leads to a 'developmental polymorphism'³⁵. Determination of the evolutionary significance of this whole organism phenomenon and its underlying molecular mechanisms should help resolve

Table 2 Mean hatching times (days) for offspring from $LDH-B^aB^a \times LDH-B^aB^a$ and $LDH-B^bB^b \times LDH-B^bB^b$ 'mass' crosses

Replicate	Offspring genotype*			
	$LDH-B^aB^a$	$LDH-B^aB^b$	$LDH-B^bB^b$	
1	11.5	(2,782)	13.0	(2,511)
2	12.5	(2,826)	13.8	(2,654)
3	12.7	(3,273)	14.3	(2,716)

Sperm and eggs from ~40 males and 37–56 females were pooled in separate dishes and then mixed to make each cross. Numbers in parentheses are the total numbers of offspring in each cross.

* The data were analysed by a single classification analysis of variance followed by Duncan's multiple range test³⁶. The three $LDH-B^aB^a \times LDH-B^aB^a$ crosses were significantly different from the three $LDH-B^bB^b \times LDH-B^bB^b$ crosses ($P < 0.05$). There were no significant differences between similar crosses.

the controversy between exponents of the 'selectionist' and 'neutralist' schools of thought.

We thank Dr Allyn Kimble for statistical advice, Judith A. DiMichele for technical assistance and Dianne Powers for artistic assistance. This work was supported by NSF grant DEB-79-12216 to D.A.P. and a NIH postdoctoral fellowship (5-F32-GM7889-02) to L.D. This is contribution 1137 from the Department of Biology.

Received 25 August 1981; accepted 24 February 1982.

- Lewontin, R. C. *The Genetic Basis of Evolutionary Change* (Columbia University Press, 1974).
- Koehn, R. K. *Science* **163**, 943–944 (1969).
- Merritt, R. B. *Am. Nat.* **106**, 173–185 (1972).
- Powers, D. A. & Powers, D. W. in *The Isozymes* Vol. 4 (ed. Markert, C.) 63–84 (Academic, New York, 1975).
- Clarke, B. *Genetics* **79**, 101–113 (1975).
- Watt, W. B. *Genetics* **87**, 177–183 (1977).
- Place, A. R. & Powers, D. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2354–2358 (1979).
- Powers, D. A., Greaney, G. S. & Place, A. R. *Nature* **277**, 240–241 (1979).
- Hoffman, R. J. *Biochem. Genet.* **19**, 129–144 (1980).
- Koehn, R. K. & Immerman, F. W. *Biochem. Genet.* **19**, 1115–1142 (1981).
- Koehn, R. K. & Siebenaller, J. F. *Biochem. Genet.* **19**, 1143–1162 (1981).
- Babalola, A. O. B., Beeststone, J. G. & Luzzatto, L. *J. biol. Chem.* **251**, 2993–3002 (1976).
- Miller, S. R., Percy, W. & Berger, E. *Biochem. Genet.* **13**, 175–188 (1975).
- Day, T. H., Hillier, P. C. & Clarke, B. *Biochem. Genet.* **11**, 141–153 (1974).
- Day, T. H. & Needham, L. *Biochem. Genet.* **11**, 167–175 (1974).
- Mitton, J. B. & Koehn, R. K. *Genetics* **79**, 97–112 (1975).
- Mitton, J. B. *Nature* **273**, 661–662 (1978).
- Johnson, M. S. *Heredity* **27**, 205–226 (1971).
- Tsuyuki, H. & Willisroft, S. N. *J. Fish. Res. Bd Can.* **34**, 996–1003 (1977).
- Redding, J. M. & Schrech, C. B. *J. Fish. Res. Bd Can.* **36**, 544–551 (1979).
- Allison, A. C. *Cold Spring Harb. Symp. quant. Biol.* **20**, 239–255 (1955).
- Johnson, F. M. & Powell, A. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1783–1784 (1974).
- Koehn, R. K., Newell, R. I. E. & Immermann, F. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5385–5389 (1980).
- Taylor, M. H., DiMichele, L. & Leach, G. J. *Copeia* 397–399 (1977).
- Whitt, G. S. *Science* **166**, 1156 (1969).
- Place, A. R. & Powers, D. A. *Biochem. Genet.* **16**, 577–591 (1978).
- Powers, D. A. & Place, A. R. *Biochem. Genet.* **16**, 593–607 (1978).
- Meid, P. A. & Powers, D. A. *J. biol. Chem.* **253**, 3521–3528 (1978).
- Powers, D. A. *Am. Zool.* **20**, 139–162 (1980).
- DiMichele, L. & Taylor, M. H. *J. exp. Zool.* **214**, 181–187 (1980).
- DiMichele, L. & Taylor, M. H. *J. exp. Zool.* **217**, 73–79 (1981).
- Able, K. W. & Castagna, M. *Chesapeake Sci.* **16**, 282–284 (1975).
- Percy, W. G. & Richards, S. W. *Ecology* **43**, 248–259 (1962).
- Meredith, W. H. & Lotrich, V. A. *Estuar. Coast. mar. Sci.* **8**, 99–118 (1979).
- Wourms, J. P. *J. exp. Zool.* **182**, 389–414 (1972).
- Ferguson, G. H. *Statistical Analysis in Psychology and Education* (McGraw-Hill, New York, 1976).

Local application of retinoic acid to the limb bud mimics the action of the polarizing region

C. Tickle*, B. Alberts†, L. Wolpert* & J. Lee*

* Department of Biology as Applied to Medicine, The Middlesex Hospital Medical School, London W1P 6DB, UK

† Department of Biochemistry, University of California, San Francisco, California 94143, USA

The polarizing region, a small group of cells at the posterior margin of the limb bud, acts as a signalling region to specify the pattern of structures which develop across the antero-posterior axis of the limb¹. When a polarizing region is grafted to the anterior margin of a second limb, a mirror-image symmetrical limb develops^{1,2}. There is indirect evidence that the signal from the polarizing region is a diffusible morphogen³. In an attempt to identify the nature of the putative morphogen, we have developed a method whereby extracts of polarizing region cells and chemicals can be locally applied to the wing bud by being bound to implanted beads and other inert carriers. The chemicals tested included hyaluronidase, dibutyl cyclic AMP and thalidomide. All experiments have been negative until, at the suggestion of Dr J. Pitts, we tested retinoic acid because of its effects on cell to cell communication and cell differentiation, and we have now for the first time been able to mimic the action of the polarizing region with a defined chemical.



Fig. 1 Dorsal view of part of a fixed embryo to show how the grafts were performed. A strip of retinoic acid-impregnated paper has been placed in an anterior position in the right wing bud.

Small strips of Whatman diethylaminoethyl cellulose (DEAE) paper (~500 μm long and 100 μm wide) were soaked in a range of concentrations of all *trans*-retinoic acid (Sigma, type XX) dissolved in dimethyl sulphoxide (DMSO; Sigma, grade 1) for 15–30 min. These strips were then grafted wet to slits cut in the anterior margins of wing buds at stages of development before the digits are specified (Hamilton–Hamburger stages 20–21). The slits to which the papers were grafted were cut into the wing at a level opposite somite 16 or the border between somites 16 and 17. Figure 1 shows the type of operation but in this case the paper has been grafted more posteriorly. The embryos were left to develop for a further 6–7 days before they were fixed and the cartilage pattern of the grafted wings examined. The results obtained are presented in Table 1.

DEAE-cellulose papers alone, papers soaked in DMSO (unpublished results of J. McQueen) and papers soaked in the lower concentrations of retinoic acid tested (0.01–2.5 mg ml^{-1}) did not produce additional digits (Fig. 2a). Instead, the most anterior digit, digit 2 and/or the anterior forearm element, the radius, is occasionally absent (digit 2 missing in 2/27 cases, the radius missing in 6/27 cases). However, papers soaked in concentrations of retinoic acid higher than 2.5 mg ml^{-1} can produce additional digits in the wings of the surviving embryos. We have previously shown that grafts of cells with high polarizing activity lead to the development of a complete duplicate set of digits: 4 3 2, in mirror-image symmetry with the normal set, 2 3 4, while grafts with progressively lower activities specify a duplicate 3 and 2 and then a duplicate 2 only^{4,5}. The pattern of duplicated digits obtained with retinoic acid suggests that the response to this chemical is similarly dose dependent. With

grafts of papers soaked in 5 mg ml^{-1} retinoic acid, 50% of the wings that develop have duplicate digits and in most cases consist of an additional digit 2 only (Fig. 2b), while with grafts of papers soaked in 7.5 mg ml^{-1} retinoic acid, nearly every wing that develops has additional digits and furthermore, half of these have duplicated both digits 4 and 3 (Fig. 2c). Duplications were also obtained at concentrations of 10 mg ml^{-1} using pure retinoic acid (Hoffmann-La Roche). Although the pattern of the production of additional digits mimics the action of the polarizing region, the detailed patterns of cartilage in the duplicated wings, in some cases, at the higher concentrations of retinoic acid resemble those typically produced when signalling is unimpaired but growth is reduced⁶.

Our assay for polarizing region activity involves placing the graft at the anterior margin of the limb bud. However, the effect of a normal polarizing region graft is dependent on its position along the antero-posterior axis. When placed at the posterior margin it has no effect. We therefore grafted retinoic acid-impregnated papers (10 mg ml^{-1}) to the posterior margin of nine wing buds. Seven gave normal wings, one lacked digits and another gave a tiny extra sliver of cartilage anteriorly. These results are consistent with local application mimicking the polarizing region rather than a systemic effect. Note that the survival of the embryos receiving grafts of the papers soaked in the high concentrations of retinoic acid that lead to the production of additional digits can be reduced (Table 1). Indeed, using two batches of retinoic acid, scarcely any embryos survived that received grafts of papers soaked in retinoic acid at 10 mg ml^{-1} , while with a third batch 100% of the treated embryos survived. Associated with the production of additional digits, embryos that had received grafts of retinoic acid-impregnated paper to the limb bud also showed beak deformities. The beak deformities, which will be described more fully elsewhere, affected the upper beak only and ranged from a complete absence of the upper beak to a slight shortening.

Retinoic acid and vitamin A have a wide variety of effects on cells, for example, inducing differentiation⁷, blocking gap junctions⁸ and affecting glycoprotein synthesis⁹. Added systemically, retinoic acid also has teratological effects on limbs, such as damaging cartilaginous elements¹⁰. It probably acts through a retinoic acid binding protein in the cytoplasm⁹.

Further work is required to determine whether retinoic acid is itself a morphogen or is converted in the limb to such a morphogen; alternatively, it may activate some other system or simply be acting as a weak acid. Moreover, a cautionary note is struck by the fact that the concentration of retinoic acid required to produce extra digits can approach that lethal to the embryo. However, cells soaked in several other toxic chemicals never produced extra digits when grafted to anterior positions in early buds¹¹. These results with retinoic acid also contrast with those studies on neutral induction, where substances producing 'sub-cytolytic' changes trigger neural differentiation without a change in pattern¹². Here, we have the first example

Table 1 The digit pattern in wings that developed following grafts of paper soaked in retinoic acid solutions to anterior positions in early wing buds

Concentration of retinoic acid solutions in which papers were soaked (mg ml^{-1})	No. of grafts	No. of survivors	Digit pattern					Other
			Normal 2 3 4	2 2 3 4	3 3 4 or 3 2 3 4	4 3 3 4 or 4 3 4	4 3 2 3 4	
0 (controls)	6	5	4	–	–	–	–	One 3 4
0.01	6	5	5	–	–	–	–	
0.1	6	6	6	–	–	–	–	
1	6	5	4	–	–	–	–	One 3 4
2.5	10	8	8	–	–	–	–	
5	31	18	9	7	1	1	–	
7.5	16	9	1	3	–	4	–	One 4
10	15	8	1	1	2	1	2	One no digits
Saturated solution	5	0						

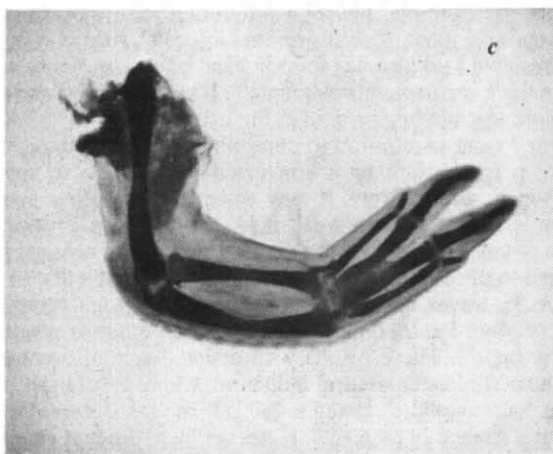
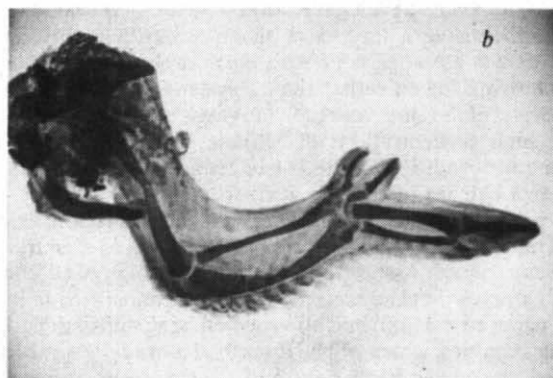
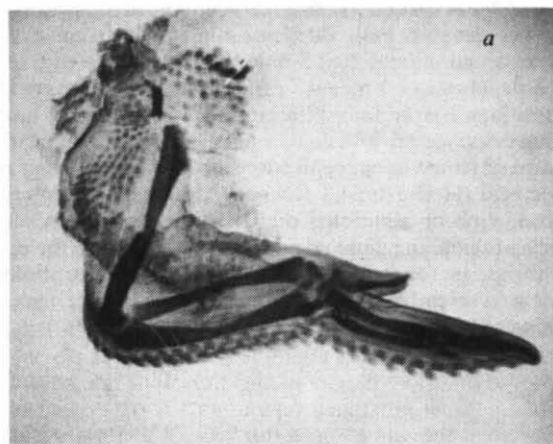


Fig. 2 Dorsal views of whole mounts of wings that developed following grafts of retinoic acid-impregnated strips of paper to anterior positions in the early wing buds. *a*, Wing that developed following a graft of a strip of paper soaked in 0.1 mg ml^{-1} retinoic acid. The digit pattern is normal: 2 3 4. Note that the strip of paper can be seen in this wing and has ended up at the head of the humerus. *b*, Wing that developed following a graft of a strip of paper soaked in 7.5 mg ml^{-1} retinoic acid. Digit pattern is 2 2 3 4. *c*, Wing that developed following a graft of a strip of paper soaked in 7.5 mg ml^{-1} retinoic acid. Digit pattern is 4 3 3 4.

of a well defined spatial pattern of cellular differentiation being created by a chemical changing positional values in a manner that closely mimics that of the normal signalling tissue.

This work was supported by the MRC and by NIH grant GM23928 to B.A. We thank A. Crawley for help with photography.

Received 30 October 1981; accepted 5 March 1982.

1. Tickle, C., Summerbell, D. & Wolpert, L. *Nature* **254**, 199–202 (1975).
2. Saunders, J. W. & Gasseling, M. T. in *Epithelial-Mesenchymal Interactions* (eds Fleischmajer, R. & Billingham, R. E.) 78–97 (Williams & Wilkins, Baltimore, 1968).
3. Tickle, C. in *Development in Mammals* Vol. 4 (ed. Johnson, M. H.) 101–137 (North-Holland, Amsterdam, 1980).
4. Smith, J. C., Tickle, C. & Wolpert, L. *Nature* **272**, 612–613 (1978).
5. Tickle, C. *Nature* **289**, 295–298 (1981).
6. Smith, J. C. & Wolpert, L. *J. Embryol. exp. Morph.* **63**, 127–144 (1981).
7. Strickland, S. & Mahdavi, V. *Cell* **15**, 393–403 (1978).
8. Pitts, J. D., Burk, R. R. & Murphy, J. P. *Cell Biol. int. Rep. Suppl.* **A5**, 45 (1981).
9. Lotan, R. *Biochim. biophys. Acta* **605**, 33–93 (1980).
10. Kochhar, D. M. *Teratology* **7**, 289–298 (1973).
11. Honig, L. S., Smith, J. C., Hornbruch, A. & Wolpert, L. *J. Embryol. exp. Morph.* **62**, 203–216 (1981).
12. Saxen, L. & Toivonen, S. *Primary Embryonic Induction*, 1–271 (Logos, London, 1962).

Long-term complications of virus-induced diabetes mellitus in mice

Ji-Won Yoon, Merlyn M. Rodrigues*, Carol Currier* & Abner Louis Notkins

Laboratory of Oral Medicine, National Institute of Dental Research and * Clinical Branch, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20205, USA

In man, the early metabolic abnormalities associated with diabetes mellitus are followed by various long-term complications¹. The cause of these complications is not clear, but they may be secondary to persistent abnormalities in blood glucose levels². In mice, several viruses can infect and destroy pancreatic β cells and produce acute-onset insulin-dependent diabetes mellitus (IDDM) characterized by hypoinsulinaemia, hyperglycaemia, glycosuria, polydipsia and polyphagia³. The virus most studied is the M variant of encephalomyocarditis (EMC) virus^{4,5}. Long-term complications, however, have been difficult to demonstrate because in most mice, the diabetes is transient and mild. Recently, by plaque purification, we have shown that the M variant of EMC virus actually has two stable variants: one diabetogenic (designated D) and the other non-diabetogenic (designated B). Further studies revealed that the B variant reduces the severity of the diabetes produced by the D variant⁶. We now report that when mice are inoculated with the D variant alone, many develop severe and prolonged hyperglycaemia followed by some of the long-term complications of diabetes, including reduced lifespan, glomerulosclerosis and ocular changes.

SJL/J male mice, 5–6 weeks old, were infected with 5×10^5 plaque-forming units (PFU) of the D variant of EMC virus and at various intervals over 6 months, non-fasting blood glucose levels were determined and mortality calculated. As seen in Table 1, 7 days after infection, 85% of the mice were diabetic with a mean blood glucose level of $426 \pm 94 \text{ mg \%}$. The number of diabetic mice decreased with time and at 150 days, only 12 of the mice were still diabetic with a mean glucose level of $290 \pm 44 \text{ mg \%}$. In contrast, age-matched uninfected controls showed no elevation of blood glucose levels. The EMC-infected diabetic animals also showed substantially higher mortality than age-matched uninfected controls or infected but non-diabetic mice (Table 1). Between 7 and 30 days after infection, only 7.1% of the diabetic mice died. At each interval thereafter, there was an increase in mortality associated with the duration of diabetes. In contrast, the infected non-diabetic mice, including those that converted from the diabetic to the non-diabetic state, showed low mortality. After 6 months, the cumulative mortality of the infected diabetic mice and uninfected controls was 64% and 17% respectively.

Evidence that the elevated blood glucose levels were responsible for the increased mortality is provided by the fact that, at

each of the intervals up to 120 days, the overall mortality of the diabetic mice having blood glucose levels above the median was about three times greater than that of those having blood glucose levels below the median.

In addition to the increased mortality, the diabetic mice showed evidence of long-term complications. At 2 months after infection, light microscopy revealed little or no change in the kidney, but after 4 months, mild to moderate mesangial thickening was found. At 6 months, the changes were quite dramatic: all 12 diabetic mice that were killed at 6 months revealed thickening of the Bowman's capsule, prominent nodular and diffuse glomerulosclerosis (Fig. 1*b*), an increase in the mesangial matrix and thickening of the tubular and peripheral glomerular basement membranes. Occasional glomeruli had capsular drop lesions, and the afferent and efferent arterioles showed evidence of hyaline sclerosis. In contrast, the glomeruli, tubules and arterioles of 12 age-matched uninfected controls (Fig. 1*a*) and 3 age-matched infected but non-diabetic mice appeared essentially normal.

Transmission electron microscopy of kidneys from diabetic mice revealed moderate thickening of the peripheral glomerular basement membrane. Approximately 20 glomeruli from each of 12 control and 12 diabetic mice (at 6 months) were examined and 5–20 measurements were made per sample as described previously⁷. The peripheral glomerular basement membrane was 91 ± 3.1 nm thick in uninfected controls (Fig. 1*c*) and $382 \pm$

6.2 nm in mice that were diabetic for 6 months (Fig. 1*d*). The foot processes were fused in some areas but intact in others, and the mesangium was thickened markedly. Uninfected age-matched controls showed normal glomerular structure, normal basement membrane and distinct delineation of the glomerular foot processes.

Approximately 10 glomeruli from each of 12 control and 12 diabetic mice (at 6 months) were examined also by scanning electron microscopy. The glomeruli of uninfected mice revealed the normal contour of Bowman's capsule, epithelial cells and foot processes (Fig. 1*e*), but in diabetic mice, there was effacement of the surface contours (Fig. 1*f*) and considerable atrophy of the glomeruli, which in some cases occupied only one-third of the Bowman's capsule.

In addition, long-term ocular abnormalities were observed in the same 12 diabetic mice that were examined for kidney lesions. The corneal epithelium of mice that were diabetic for 6 months showed marked irregularity of all layers (Fig. 2*b*) compared with age-matched uninfected controls (Fig. 2*a*). In particular, the basal layer and the junctional complexes were focally deficient. The stroma and Descemet's membrane were normal. We prepared retinal capillaries from 6–12 diabetic and control mice by trypsin digestion⁸ and examined them by light and scanning electron microscopy. Capillaries from mice that were diabetic for 6 months revealed a moderate decrease (~50%) in the number of pericytes with the loss of the usual

Fig. 1 Kidney sections from uninfected controls and EMC-infected mice that were diabetic for 6 months. Light microscopy of *a*, control mouse showing normal glomeruli, tubules and Bowman's capsule (double arrows); *b*, diabetic mouse (6 months duration) showing prominent nodular glomerulosclerosis (long arrows) and thickening of Bowman's capsule (double arrows) (periodic acid-Schiff, $\times 145$). Transmission electron micrographs of kidney from a diabetic mouse (*d*) showing marked thickening of the peripheral glomerular capillary basement membrane (I—I) compared with an uninfected control (*c*) ($\times 15,750$). Scanning electron micrograph of kidney from diabetic mouse showing *f*, focal effacement of the normal glomerular epithelial cells with loss of foot processes (arrow), compared with *e*, the usual surface pattern of epithelial cells having abundant arborizing foot processes (double arrow) in normal kidney ($\times 3,700$). Sections were prepared by standard methods and examined using a Phillips 400 electron microscope or a JEOL SM 35 scanning electron microscope.

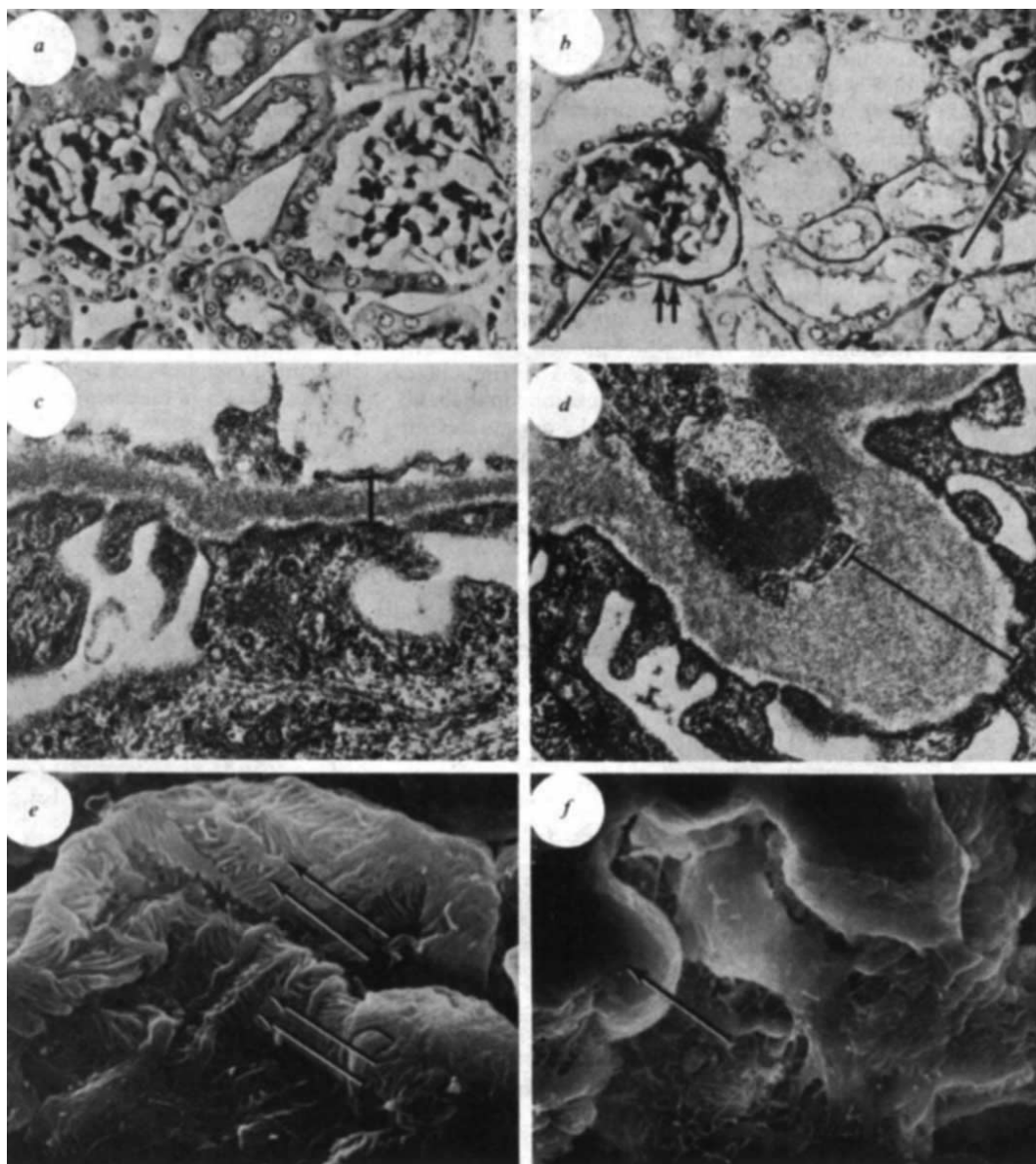


Table 1 Blood glucose levels and mortality of EMC-infected diabetic and non-diabetic mice

Days after infection	No. of animals at start of each period	Infected								
		Diabetic			Non-diabetic			Uninfected controls		
		Glucose	Deaths/ diabetics	Mortality (%)	Glucose	Deaths/ diabetics	Mortality (%)	Glucose	Deaths/ controls	Mortality (%)
7-30	100	426±94	6/85	7.1	174±29	0/15	0	156±5	0/89	0
30-60	94	401±92	16/78	20.5	166±19	0/16	0	152±14	1/89	1.1
60-90	79	367±81	10/61	16.4	163±13	0/18	0	152±11	4/88	4.6
90-120	68	332±77	16/43	37.2	159±14	0/25	0	154±12	5/85	5.9
120-150	52	306±70	9/21	42.9	167±20	0/31	0	159±16	2/80	2.5
150-180	43	290±44	7/12	58.3	168±23	1/31	3.2	162±19	3/78	3.9

A mouse was scored as diabetic if the non-fasting glucose level was ≥ 3 s.d. above the mean of uninfected controls. The glucose concentrations shown are the mean $\pm$ s.d. at the beginning of each period; those shown for the infected non-diabetic animals are < 3 s.d. above the mean of uninfected mice. Fourth column, number of deaths at end of each period divided by number of diabetic mice at beginning of each period. Column 7, number of deaths at end of each period divided by number of non-diabetic mice at beginning of each period. Column 10, number of deaths at end of each period divided by number of uninfected controls at beginning of each period.

1:1 ratio of pericytes to endothelial cells (Fig. 2c,d). In ~60% of the diabetic mice, the retinal capillaries showed a minimal to moderate (twofold) increase in the thickness of the basement membranes. Microaneurysms were not detected in any of the animals by the PAS-trypsin digestion method^{8,9} and no cataracts were detected. As previously observed¹⁰, the pancreatic tissues showed marked atrophy of β cells in the islets of Langerhans and preservation of acinar tissue.

Thus we have shown that the decrease in the mean blood glucose level of EMC-infected mice over the 6 months of the experiment (Table 1) was due to several factors. First, many of the highly diabetic mice died, leaving survivors having much lower blood glucose levels. Second, monthly monitoring of individual animals revealed that the blood glucose of these animals also decreased with time (data not shown). Moreover, ~20% of the originally diabetic animals converted to the non-diabetic state by 150 days after infection. Whether regeneration of β cells or other compensatory factors is responsible for the decrease in the severity of the diabetes with time is unknown.

Using the M variant of EMC virus, which produces less severe diabetes, and by examining mice 4 months after infection, Kanich *et al.*¹¹ observed some mesangial thickening in diabetic mice by transmission electron microscopy. In our studies on the D variant of EMC virus, not only did we find mesangial

thickening, but also an increase in the thickness of the peripheral glomerular basement membrane and Bowman's capsule. In addition, we found prominent nodular and diffuse glomerulosclerosis, as well as afferent and efferent arteriosclerosis. These changes are compatible with the renal complications seen in human diabetes mellitus, especially of the Kimmelstiel-Wilson type¹². The most striking similarity to human diabetic nephropathy was the change in the peripheral glomerular basement membrane, which was thickened by as much as fourfold compared with control mice. Such changes have not been well documented in alloxan-induced or streptozotocin-induced diabetes, which generally have been associated with varying degrees of mesangial thickening without significant thickening of the peripheral glomerular basement membranes^{9,13}.

The histopathological changes seen in the cornea of our diabetic mice were not dissimilar to those which predispose to corneal erosion in human diabetes¹⁴. Moreover, the decrease in the number of pericytes in the retinal vessels of diabetic animals, with the loss of the usual 1:1 ratio of pericytes to endothelial cells and the moderate increase in thickness of the basement membrane of retinal capillaries, is also characteristic of early stages of diabetes in man¹⁵. These changes are thought to precede microaneurysms. Microaneurysms usually occur in long-term diabetes in humans and in diabetic animals such as

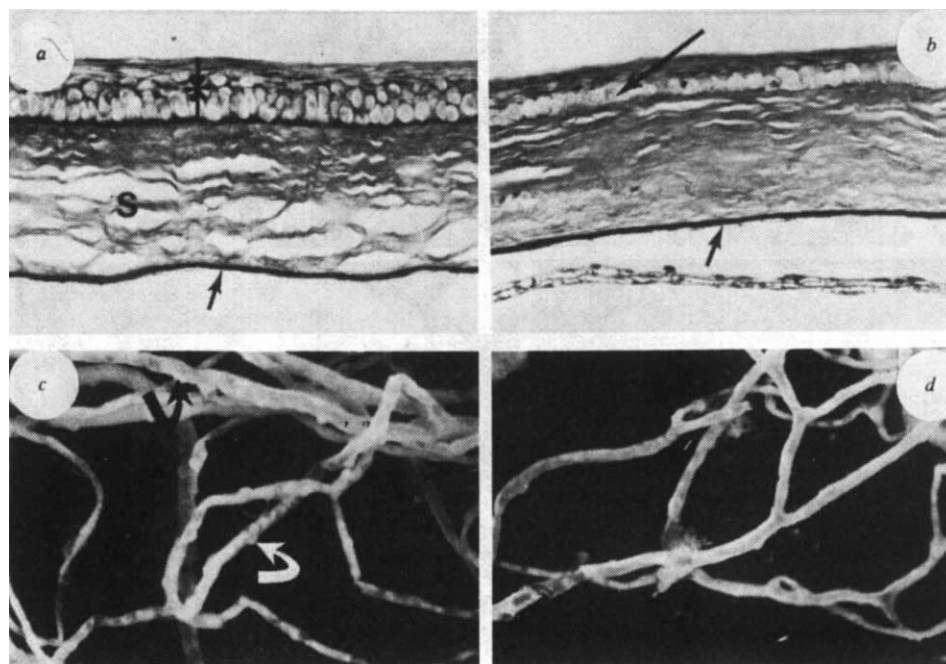


Fig. 2 The corneal epithelium of uninfected mice (a) show the usual stratification (cross) and glycogen content (periodic acid-Schiff; $\times 236$). Cornea from mice that were diabetic for 6 months (b) show irregular stratification and moderate oedema of the epithelium, particularly of the basal layer (long arrow) (periodic acid-Schiff; $\times 230$). The stroma (S) and Descemet's membrane (short arrow) are normal. c, Scanning electron micrograph of retinal capillaries obtained by trypsin digestion from uninfected mice, showing the normal pericyte distribution (arrows). d, A moderate decrease in the number of pericytes is observed in the trypsin digest of retinal capillaries from mice that were diabetic for 6 months ($\times 308$).

the monkey and dog, but have not been demonstrated convincingly in small animals such as the mouse and rat⁸. The possibility that microaneurysms may develop in EMC-infected mice that are diabetic for >6 months is now being investigated.

In conclusion, our studies in mice show that infection with the D variant of EMC virus produces both the acute and at least some of the long-term complications of IDDM.

Received 26 October 1981; accepted 12 February 1982.

1. Rep. of the National Commission on Diabetes to the Congress of the United States (U.S. Department of Health, Education and Welfare, Government Printing Office, 1976).
2. Cahill, G. F., Etzwiler, D. D. & Freinkel, N. *New Engl. J. Med.* **294**, 1004–1005 (1976).
3. Notkins, A. L. *Scient. Am.* **241**(5), 62–73 (1979).
4. Craighead, J. E. *Prog. med. Virol.* **19**, 161–214 (1975).
5. Yoon, J. W. & Notkins, A. L. *Postgrad. Med.* (in the press).
6. Yoon, J. W., McClintock, P. R., Onodera, T. & Notkins, A. L. *J. exp. Med.* **152**, 878–892 (1980).
7. Siperstein, M. *Adv. intern. Med.* **18**, 325–344 (1972).
8. Cogan, D. G., Toussaint, D. & Kuwabara, T. *Arch. Ophthalmol.*, N.Y. **66**, 366–378 (1961).
9. Engerman, R. L. *Trans. Am. Acad. Ophthalmol. Otolaryng.* **81**, 710–715 (1976).
10. Hayashi, K., Boucher, D. W. & Notkins, A. L. *Am. J. Path.* **75**, 91–102 (1974).
11. Kanich, R. E., Craighead, J. E. & Kessler, J. B. *Diabetologia* **9**, 203–209 (1973).
12. Heptinstall, R. H. in *Pathology of the Kidney* 2nd edn, 929–974 (Little, Brown & Co., London, 1974).
13. Steen-Olsen, T. *Int. Rev. exp. Path.* **7**, 271–304 (1969).
14. Taylor, H. R. & Kinsey, R. A. *Invest. Ophthalmol. vis. Sci.* **20**, 548–553 (1981).
15. Ashton, N. *Br. J. Ophthalmol.* **58**, 344–366 (1974).

Neural regulation of muscarinic receptors in chick expansor secundariorum muscle

R. A. Rush, M. F. Crouch, C. P. Morris & B. J. Gannon

Centre for Neuroscience and Departments of Human Physiology and Human Morphology, The Flinders University of South Australia, Bedford Park, South Australia 5042, Australia

We have recently observed the loss of muscarinic receptors from a smooth muscle of the chicken wing, the expansor secundariorum, during post-hatch development¹. The nerve supplying the muscle has been shown to consist mainly of sympathetic fibres that, when stimulated, cause muscle contraction by the release of noradrenaline and subsequent activation of α -adrenoreceptors. Although no acetylcholine release has been detected in isolated nerve-muscle preparations², the presence of a cholinergic input has not been clearly excluded. Kuromi and Hagihara² observed that atropine was ineffective in blocking the nerve stimulation-induced contraction of the muscle, whereas phentolamine partly and guanethidine totally abolished the contraction. We have now examined the effect of denervation of the muscle 1 day after hatch on the binding of ³H-quinuclidinyl benzilate (³H-QNB), a radioligand for muscarinic receptors^{3,4}. Our results indicate that the normal loss of muscarinic receptors from the expansor secundariorum is prevented by denervation. This observation is important as it demonstrates the regulation of a receptor population by a nerve supply that apparently does not release the corresponding transmitter.

Muscles from 1-day old birds bound 0.165 ± 0.024 fmol of ³H-QNB per mg wet weight. In 3-week old chicks that were unilaterally denervated at 1 day old, no significant loss of binding sites occurred (0.140 ± 0.018 fmol per mg wet weight, $P > 0.1$), whereas binding in the contralateral, innervated muscle was significantly lower (0.093 ± 0.013 , fmol per mg wet

weight, $P < 0.001$). An anomalous, small difference in the binding was detected between the contralateral, innervated muscle and muscle from normal, unoperated birds (Fig. 1).

The weight of this muscle increased after denervation⁵, the average weight of the denervated muscles being 15.8 ± 1.5 mg, which was 48% greater than for innervated muscles (10.7 ± 1.0 mg; $P < 0.001$, $n = 32$). Consequently, the difference in ³H-QNB binding between the innervated and denervated muscles was exaggerated when expressed as amount bound per whole muscle. The total ³H-QNB bound by the denervated muscles was 1.98 ± 0.15 fmol per muscle and this was 2.1 times the amount bound by the innervated muscle (0.90 ± 0.08 fmol per muscle).

The effectiveness of sympathetic denervation procedures was verified in several animals by fluorescence histochemistry⁶. Stretch preparations of glyoxylic acid-treated muscle showed a dense innervation of the normal muscle but almost complete absence of fibres in the denervated tissue. Occasionally a few regenerating fibres were observed at the edge of the muscle.

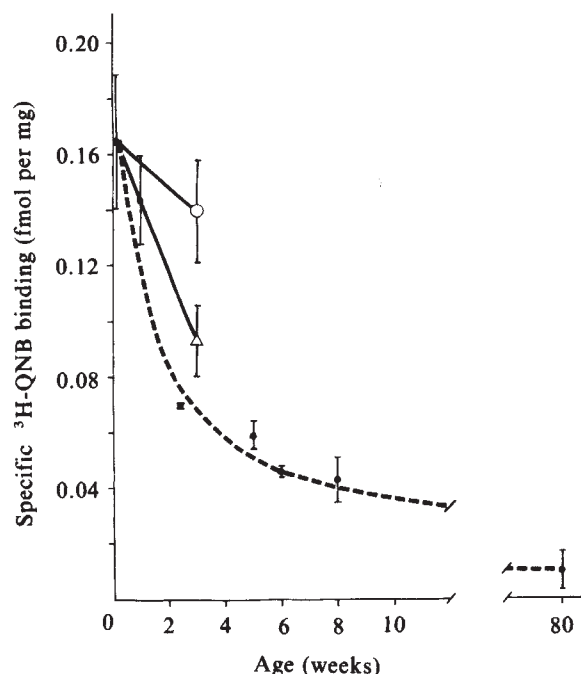


Fig. 1 Effect of denervation on the muscarinic receptor population of the developing chicken secundariorum muscle. The muscarinic receptor population of unoperated chicks at various ages (broken lines) is compared with that of chicks denervated unilaterally at 1 day old and killed 3 weeks later (solid lines). The open circles indicate the ³H-QNB bound by denervated muscles, and the open triangles the contralateral, control muscles. Denervation of the muscle was performed unilaterally by cauterization of the nerve 1 cm from its entry into the muscle. Three weeks after denervation, muscles were excised and homogenized with a glass-glass hand homogenizer (Wheaton Scientific) in 1.5 ml of 50 mM phosphate buffer, pH 7.4. After centrifugation at 18,000g for 4 min, pellets were resuspended in 5 or 10 volumes of buffer. Triplicate 200- μ l samples were incubated at 37°C for 1 h with ³H-QNB (1 nM, NEN) in a total volume of 240 μ l. Incubation was terminated with the addition of 1 ml of buffer and centrifugation for 4 min at 18,000g. Following two buffer washings, pellets were resuspended in 200 μ l of distilled water, added to 15 ml Triton X-100/toluene scintillation cocktail and counted in a Searle Nucleonics β -spectrometer at an efficiency of ~40%. As 10–20 mg wet weight of original muscle were required for each determination, it proved necessary to pool varying numbers of muscles. Each point represents a minimum of 11 individual determinations, and is expressed as the mean $\pm$ s.e.m. All statistical values were obtained using a two-tailed Student's *t*-test.

Several radioligands have been used to estimate muscarinic binding sites in brain⁷, sympathetic ganglia⁸ and smooth muscle^{3,4}. These studies have indicated that the use of radio-labelled muscarinic antagonists and in particular ³H-quinuclidinyl benzylate results in binding to a single site that is accepted as the muscarinic receptor. Our previous studies with the expensor secundariorum have shown that the binding characteristics of the radioligand in this tissue are similar to those observed in other tissues, saturation occurring at 0.6 nM and with a K_D of 0.089 nM⁻¹. We have therefore used the binding of ³H-QNB to muscle membrane as a measure of the muscarinic receptor population.

Trophic regulation of skeletal muscle function has long been the subject of intensive investigation^{9,10}. In contrast, few studies have examined the trophic effect of autonomic innervation on smooth muscle function. Indeed, until recently there has been little evidence that postganglionic sympathetic fibres exert a trophic influence on their effector cells. Unlike skeletal muscle, smooth muscle demonstrates a remarkable degree of autonomy. In particular, atrophy does not occur following denervation but functional integrity is retained.

Our results provide strong evidence for a trophic regulation of smooth muscle function by sympathetic nerves. Prevention of the normal loss of ³H-QNB binding sites from the developing expensor secundariorum by denervation suggests that their production is normally repressed by a mechanism of neuronal origin that is likely to be mediated by transmitter molecules, specific trophic molecules or electrical activity within the muscle. Whether the low level of muscarinic receptors in this tissue has any function in the young animal has not been determined.

The composition of the nerve supplying the expensor secundariorum has not been fully established. Although it is clear that there is a dense adrenergic innervation of the muscle¹¹, no information has been presented concerning a sensory input. Nerve stimulation has been shown to cause muscle contraction via noradrenaline release, with no evidence of acetylcholine release². These studies clearly show that the nerve regulates the muscarinic receptor population, but do not demonstrate whether this regulation is mediated by noradrenergic fibres, sensory fibres or other as yet unidentified nerve types. The evidence against the involvement of cholinergic nerves is not conclusive, but makes it unlikely that acetylcholine is involved. Histochemical and biochemical analyses over the entire developmental period are necessary to exclude this possibility. Further studies examining the origin of the nerve supply to this muscle would also aid in assessing the factor or factors involved in the regulation of the muscarinic receptor population.

Finally, the presence of a discrete nerve supply and tendon to the expensor secundariorum, together with the results of the present study, suggests that this unique muscle will be a useful tissue in which to study the interaction between smooth muscle and its innervation.

This work was funded by the NH and MRC of Australia.

Received 9 October 1981; accepted 15 February 1982.

1. Crouch, M. F., Morris, C. P., Cheah, T. B. & Rush, R. A. *Brain Res.* **232**, 212–215 (1982).
2. Kuromi, H. *Eur. J. Pharmac.* **51**, 285–289 (1978).
3. Yamamura, H. I. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1725–1729 (1974).
4. Yamamura, H. I. & Snyder, S. H. *Molec. Pharmac.* **10**, 861–867 (1974).
5. Campbell, G. R., Allan, B. I. & Gannon, B. J. *Cell Tissue Res.* **176**, 143–156 (1977).
6. Axelsson, S., Bjorkland, A., Falck, B., Lindwall, O. & Svensson, L. A. *Acta physiol. scand.* **87**, 57–62 (1973).
7. Hulme, E. C., Birdsall, N. J. M., Burgen, A. S. V. & Mehta, P. *Molec. Pharmac.* **14**, 737–750 (1978).
8. Birt, D. R. *Brain Res.* **143**, 573–579 (1978).
9. Guth, L. *Physiol. Rev.* **48**, 645–687 (1969).
10. Harris, A. J. in *Physiology of Peripheral Nerve Disease* (ed. Sumner, A. J.) 195–220 (Saunders, London, 1980).
11. Bennett, T. E. & Malmfors, T. Z. *Zellforsch.* **106**, 22–50 (1970).

Chronic continuous administration of neuroleptic drugs alters cerebral dopamine receptors and increases spontaneous dopaminergic action in the striatum

K. Murugiah, A. Theodorou, S. Mann, A. Clow, P. Jenner & C. D. Marsden*

University Department of Neurology, Institute of Psychiatry and King's College Hospital Medical School, Denmark Hill, London SE5 8AF, UK and Agricultural Research Council, Institute of Animal Physiology, Babraham, Cambridge CB2 4AT, UK

The 'dopamine hypothesis' of schizophrenia states that the illness is due to overactivity of dopamine mechanisms in the brain. This hypothesis is based on two facts: (1) drugs, such as amphetamine, that enhance dopaminergic neurotransmission in the brain, may occasionally provoke a schizophrenic psychosis; and (2) acute administration of neuroleptic drugs, which are used to treat schizophrenia and other psychotic illnesses, causes blockade of brain dopamine receptors and initiates a chain of compensatory events which attempt to overcome such an action. We have previously shown that administration of neuroleptic drugs to rats for up to 18 months produces unexpected effects^{1,2}; after 6 months, all signs of blockade of dopamine receptors in the striatum have disappeared, and thereafter striatal dopamine receptors increase in number and become behaviourally supersensitive to administered dopamine agonists such as apomorphine. We now show that such chronic exposure to neuroleptics completely alters other characteristics of striatal dopamine receptors, and that in the intact animal, these changes are associated with spontaneous overactivity of striatal dopaminergic mechanisms. Unless other dopaminergic brain areas, for example in the mesolimbic or mesocortical regions, react differently from the striatum, the mode of action of neuroleptics, and the dopamine hypothesis of schizophrenia, may have to be reconsidered.

We first investigated the changes in the biochemical characteristics of striatal dopamine receptors which take place during 12 months' continuous administration of the neuroleptic trifluoperazine dihydrochloride to rats in two separate experiments—in one of them the drug was subsequently withdrawn and the animals were followed for a further 6 months (Fig. 1). The characteristics of striatal dopamine receptors were examined in these animals by estimating ³H-spiperone (0.125–4.0 nM) binding to striatal homogenates, the specificity of binding to dopamine receptors being defined by incorporation of dopamine (10^{-4} M). After the first 3 months of trifluoperazine treatment there was a progressive increase in the number of striatal dopamine receptor binding sites (B_{max}) for ³H-spiperone binding in the drug-treated animals, compared with the age-matched controls. B_{max} continued to increase until administration of the drug ceased after 1 yr; the number of striatal dopamine receptors then began to revert towards control levels after 3 months' drug withdrawal, and had reached control values after 6 months' drug withdrawal.

Receptor affinity, as judged by the dissociation constant (K_d) for ³H-spiperone binding to striatal tissue homogenates, also changed during 1 yr administration of trifluoperazine. In the first 3 months, K_d increased, as predicted because of the presence of drug in the tissue analysed, but after 6 months of drug administration K_d had completely or partially reverted to control levels. Thereafter, in both experiments, K_d progressively

* To whom correspondence should be addressed.

rose again, until the drug was withdrawn, when it abruptly reverted to control values. These changes could be explained if the drug levels had fallen, despite continued administration, perhaps due to enzyme induction, but this would not explain why K_d rose again between 6 and 12 months of trifluoperazine administration. In any case, measurement of the plasma concentration of the drug showed a gradual increase throughout this period of administration (G. Watkins, R. Whelpton and S. Curry, personal communication). The change in striatal dopamine receptor affinity after 6 months of trifluoperazine intake occurred simultaneously with the increase of dopamine receptor numbers (B_{max}), suggesting that a new population of dopamine receptors of higher affinity than normal might have appeared; but a Hill plot of the data for 3H -spiperone binding to striatal homogenates did not deviate from unity. Whatever the reason for the change in the characteristics of striatal dopamine receptors after 6 months of trifluoperazine administration, and we can offer no explanation, the change

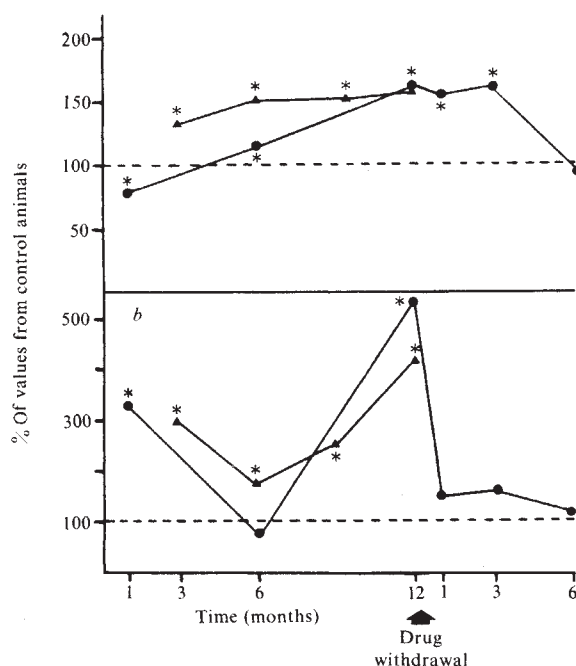


Fig. 1 Comparison of the alterations in dopamine (10^{-4} M)-specific 3H -spiperone (0.125 – 4.0 nM) binding to rat striatal preparations produced by the administration of trifluoperazine dihydrochloride continuously for up to 12 months in two distinct studies. *a*, Receptor numbers (B_{max}); *b*, dissociation constant (K_d). The experimental groups were as follows. Expt 1: ●: male Wistar rats (initial weight 200 ± 12 g; Olac International) received trifluoperazine dihydrochloride (2.5 – 3.5 mg per kg per day; Smith, Kline and French) for up to 12 months when the drug was withdrawn. Expt 2: ▲: male Wistar rats (initial weight 176 ± 4 g; Bantin and Kingman) received trifluoperazine dihydrochloride (4.5 – 5.6 mg per kg per day) for up to 12 months. In each case a group of age-matched control animals from the same supplier received distilled water alone and was maintained alongside the drug-treated animals. At intervals during the course of the experiments, striatal preparations from drug-treated and control animals were examined for dopamine (10^{-4} M)-specific 3H -spiperone (26 Ci mmol $^{-1}$; Amersham) binding, according to the method of Lysen *et al.*¹⁶. Rats from age-matched control and drug-treated groups ($n = 7$) were killed by cervical dislocation and decapitation and the brain rapidly removed and placed on ice. The paired corpora striata were removed and placed in ice-cold 50 mM Tris buffer pH 7.7 . Specific binding of 0.125 – 4.0 nM 3H -spiperone, as defined in the presence and absence of 10^{-4} M dopamine, was determined in triplicate on the pooled striatal tissue from both groups. The data obtained were subjected to Scatchard analysis to determine the number of binding sites, expressed as pmol per g wet weight of tissue (B_{max}), and the dissociation constant (K_d). Statistical analysis of the receptor binding data was carried out using linear regression analysis of the results for drug-treated and control animals, which were compared using a two-tailed Student's *t*-test. * $P < 0.05$. For comparison of the dissociation constant (K_d) and receptor numbers (B_{max}) in different animal groups at different time intervals, values in drug-treated animals are expressed as a percentage of K_d and B_{max} values in tissue from control animals obtained on the same day and assayed in parallel throughout the experiment.

Table 1 The effect of continuous administration for 14 months of distilled drinking water containing *cis*-flupenthixol (0.8 – 1.2 mg per kg per day), compared with distilled drinking water alone, on apomorphine-induced stereotyped behaviour and striatal acetylcholine concentrations in rats

Drug treatment (mg per kg per day)	Stereotypy score	Striatal acetylcholine (μ g g $^{-1}$)
Controls (distilled water alone)	2 ± 0 (6)	4.1 ± 0.3 (8)
<i>cis</i> -flupenthixol (0.8 – 1.2)	$3.9 \pm 0.1^*$ (6)	$7.4 \pm 0.7^*$ (8)

Male Wistar rats (initial weight 200 ± 12 g; Olac International) received *cis*-flupenthixol hydrochloride (0.8 – 1.2 mg per kg per day; Lundbeck) for 14 months. A group of age-matched controls receiving distilled water alone was maintained alongside the drug-treated animals. Apomorphine-induced (0.5 mg per kg subcutaneously 15 min previously; Macfarlan Smith) stereotypy was assessed according to the scale: 0 = normal; 1 = continuous locomotor activity, discontinuous sniffing; 2 = continuous sniffing; 3 = discontinuous biting, licking and gnawing; 4 = continuous biting, licking and gnawing¹². The results are expressed as the mean (± 1 s.e.m.) for six animals in each group. Acetylcholine levels were determined by bioassay. The caudate nuclei of treated and control rats were rapidly dissected from the brain and homogenized in 1% acetic ethanol (1 ml glacial acetic acid in 99 ml absolute ethanol) and stored at -70°C until required. Each sample was removed from the deep freeze and rotary evaporated to dryness. 4 ml of frog Ringer were added to the extract and the pH was adjusted to 3.0 where necessary. The extract was then assayed on the eserized frog rectus using the method of Chang and Gaddum¹³, more recently described by Hebb and Silver¹⁴. Samples of the extract were compared with standard solutions of acetylcholine. A quantity of boiled alkali-treated sample equal to that of the unknown was added with each standard¹⁵ to destroy the acetylcholine, and the addition of this as a control ensured that any substances (such as potassium) in the sample capable of causing a contraction of the rectus were not assayed as acetylcholine. To sensitize and stabilize the muscle preparation, 0.5 mg of Na-ATP was added with each sample and standard to the muscle bath. The assay used in this way was accurate to within 10%. Values in parentheses give the number of observations.

* $P < 0.05$ compared with control animals, a two-tailed Student's *t*-test was used for parametric data; Mann-Whitney *U*-test was used for non-parametric results.

coincided with the onset of a behaviourally supersensitive response by the rats to the administration of dopamine agonists.

In the second series of experiments rats received *cis*-flupenthixol hydrochloride continuously for periods of up to 18 months. These animals demonstrated the same changes in B_{max} and K_d for specific binding of 3H -spiperone to striatal preparations. K_d initially increased, then reverted towards control levels at around 6 months, only to increase again by 12–18 months of drug intake (data not shown). At 18 months, prolonged neuroleptic intake had not only caused an increase in the number of striatal dopamine receptors, but also altered the relative sensitivity of the animals to the behavioural effects of dopamine agonists and antagonists. Whereas administration of apomorphine hydrochloride (0.125 – 1.0 mg per kg subcutaneously) to animals receiving such neuroleptic treatment resulted in an enhanced stereotyped response compared with age-matched controls, administration of an acute dose of trifluoperazine dihydrochloride (1 – 8 mg per kg intraperitoneally) produced less catalepsy in *cis*-flupenthixol-treated animals than in the control group (Fig. 2). Thus after 18 months' intake of *cis*-flupenthixol, the animals appeared to possess an increased number of striatal dopamine receptors that were supersensitive to the behavioural effects of dopamine agonists, but subsensitive to the effects of dopamine antagonists.

These results indicate that chronic neuroleptic administration alters the characteristics of striatal dopamine receptors as measured *in vitro* by receptor binding techniques, or as challenged by dopamine agonists and antagonists in behavioural pharmacological experiments. The crucial issue, however, is whether such changes are functionally significant in the intact animal or human patient.

To test this possibility, we have studied the effect of chronic neuroleptic intake on striatal acetylcholine activity. The postsynaptic striatal dopamine receptor lies, at least in part, on the cell body of cholinergic neurones, on which dopamine exerts inhibitory control^{3,4}. Alteration of striatal dopaminergic function affects acetylcholine activity. The acute administration of

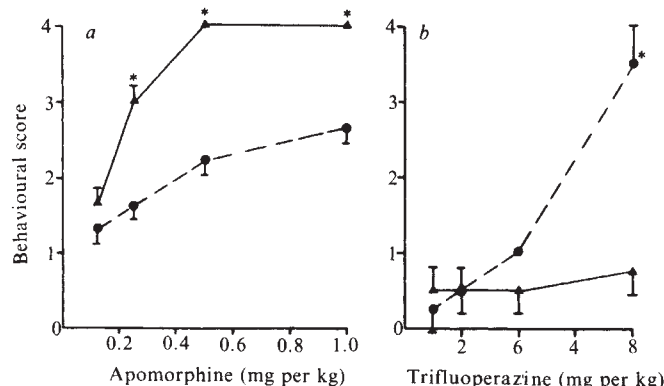


Fig. 2 Dose-response curves for *a*, the induction of apomorphine-induced stereotypy; and *b*, trifluoperazine-induced catalepsy in rats (—▲—) receiving *cis*-flupenthixol (0.8–1.2 mg per kg per day) for 18 months, compared with control animals receiving distilled water alone (---●---). Male Wistar rats (initial weight 200 ± 12 g; Olac International) received *cis*-flupenthixol hydrochloride (0.8–1.2 mg per kg per day; Lundbeck) for 18 months. A group of age-matched control animals receiving distilled water alone was maintained alongside the drug-treated animals. Stereotypy induced by apomorphine (0.125–1.0 mg per kg subcutaneously 15 min previously; Macfarlan Smith) was assessed as described in Table 1 legend. The results are expressed as the mean (± 1 s.e.m.) for six animals in each group. Catalepsy induced by trifluoperazine dihydrochloride (1–8 mg per kg intraperitoneally 3 h previously; SKF) was assessed as the time a rat maintained its front paws over a metal bar 2 cm in diameter and 7 cm from bench level and was scored according to the scale: 0–9 s = 0; 10 s–2.5 min = 1; 2.6–5 min = 2; 5.1–10 min = 3; 10.1–20 min = 4; >20 min = 5. The results are expressed as the mean (± 1 s.e.m.) for six animals in each group. * $P < 0.05$ compared with control animals using a Mann-Whitney *U*-test for non-parametric data.

Received 16 July 1981; accepted 15 February 1982.

1. Clow, A., Jenner, P., Theodorou, A. & Marsden, C. D. *Nature* **278**, 59–61 (1979).
2. Clow, A., Jenner, P. & Marsden, C. D. *Eur. J. Pharmac.* **57**, 365–375 (1979).
3. Connor, J. D. *J. Physiol., Lond.* **208**, 691–703 (1970).
4. Hattori, T., Singh, V. K., McGeer, E. G. & McGeer, P. L. *Brain Res.* **102**, 164–173 (1976).
5. Sethy, V. H. & van Woert, M. H. *Res. Commun. chem. Path. Pharmac.* **8**, 13–28 (1974).
6. Trabucchi, M., Cheney, D., Racagni, G. & Costa, E. *Nature* **249**, 664–666 (1974).
7. Stadler, H., Lloyd, K. G., Gadea-Ciria, M. & Bartholini, G. *Brain Res.* **55**, 476–480 (1973).
8. Consolo, S., Ladinsky, H. & Garattini, S. *J. Pharm. Pharmac.* **26**, 275–277 (1974).
9. Consolo, S., Ladinsky, H. & Bianchi, S. *Eur. J. Pharmac.* **33**, 345–351 (1975).
10. Consolo, S., Ladinsky, H., Bianchi, S. & Ghezzi, D. *Brain Res.* **135**, 255–263 (1977).
11. Sherman, K. A., Hanin, I. & Zigmond, J. J. *Pharmac. exp. Ther.* **206**, 677–686 (1978).
12. Costall, B. & Naylor, R. J. *Eur. J. Pharmac.* **27**, 46–58 (1974).
13. Chang, H. C. & Gaddum, J. H. *J. Physiol., Lond.* **79**, 255–285 (1933).
14. Hebb, C. O. & Silver, A. in *Neuropathology: Methods and Diagnosis* (ed. Tedeschi, C. G.) 665–690 (Little Brown, Boston, 1970).
15. Feldberg, W. & Hebb, C. J. *J. Physiol., Lond.* **106**, 8–17 (1947).
16. Laysen, J. E., Gommeron, E. & Laduron, P. M. *Biochem. Pharmac.* **27**, 307–316 (1978).

Specific non-opiate binding sites for human β -endorphin on the terminal complex of human complement

L. Schweigerer*, S. Bhakdi† & H. Teschemacher*

* Rudolf Buchheim-Institut für Pharmakologie und

† Institut für Medizinische Mikrobiologie der Justus Liebig-Universität Gießen, Frankfurter Strasse 107, D-6300 Gießen, FRG

The opioid peptide β -endorphin is released from the pituitary into the blood under physical or emotional stress; its target tissues, however, are unknown¹. We report here the specific binding of human β -endorphin to both terminal complexes of human complement: the cytolytic, membrane-derived C5b-9(m) complex and the cytolytically inactive, serum-derived SC5b-9 complex. Binding is through the carboxy-terminus of β -endorphin rather than the amino-terminus, which is implicated in binding to opiate receptors. It is conceivable that the binding of β -endorphin to complement has physiological relevance in the adaptation of the immune system to stress.

Human serum or plasma drawn in 10 mM EDTA was obtained from healthy donors. The complement pathway was activated by adding 2 g crushed inulin (Merck) to 100 ml serum followed by incubation at 37 °C for 4 h. Activation by inulin leads to the formation of the macromolecular, non-cytolytic SC5b-9 complex (comprising of complement components C5b, C6, C7, C8, C9 and the S-protein²). Supernatants obtained after centrifugation of the inulin-serum suspension were used directly in binding experiments or served as a starting material for the isolation of the SC5b-9 complex^{2,3}. The terminal, cytolytic complement complex, C5b-9(m), which consists of complement components C5b to C9 without the S-protein, was generated on target sheep erythrocyte membranes by lysis of the antibody-coated cells with human serum, and then isolated from detergent solubilizates of the membranes as described previously⁴. Detergent-solubilized, hypotonically lysed sheep erythrocyte membranes served as controls.

¹²⁵I-labelled β -endorphin was obtained by iodinating human β -endorphin (Bachem) using published techniques^{5,6}. Binding of human β -endorphin to the terminal C5b-9 complexes was demonstrated in a radioreceptor assay and by sedimentation analysis in sucrose density gradients.

In the radioreceptor assay, a small amount of human ¹²⁵I- β -endorphin was found to bind to one or several components of native human serum (Fig. 1). After inulin activation, however, binding increased considerably (Fig. 1).

When ¹²⁵I- β -endorphin was incubated with native serum or with EDTA-plasma and the samples were then centrifuged through linear sucrose density gradients, total radioactivity was recovered in the top fractions of the gradient (not shown). In

neuroleptics decreases the content of acetylcholine in the striatum, due to an increased use of acetylcholine in active cholinergic neurones^{5–11}. If chronic neuroleptic administration causes a functional reversal of initial dopamine receptor blockade in the intact animal, the striatal acetylcholine content should increase, due to a decrease in cholinergic activity resulting from enhanced dopaminergic inhibition of striatal cholinergic neurones.

We produced striatal dopamine receptor supersensitivity in rats by the continuous administration of *cis*-flupenthixol for 14 months and examined changes in striatal acetylcholine content in the same animals. At the end of 14 months treatment with *cis*-flupenthixol, the rats exhibited an enhanced behavioural response to the dopamine agonist apomorphine, and basal striatal acetylcholine content had doubled (Table 1). This result suggests that after chronic *cis*-flupenthixol treatment, not only was the response of striatal dopamine receptors to an exogenous dopamine agonist no longer antagonized, but also the receptors were overactive in the intact animal. In other words, the chronically neuroleptic-treated rat striatum exhibited spontaneous dopaminergic synaptic overactivity.

In conclusion, chronic neuroleptic intake by rats for periods of up to 12–18 months changes fundamentally the character of striatal dopamine receptors, which increase in number, but also exhibit unexplained alterations in their affinity. Such changes are associated with parallel alterations in the response of the animals to exogenous dopamine agonists and antagonists. In addition, and most important, there are also changes in striatal acetylcholine function which suggest that the spontaneous activity of at least some striatal dopamine receptors is increased in the chronically neuroleptic-treated animal. These events may be responsible for the emergence of tardive dyskinesias in man, and they must be taken into account when considering the mechanism of action of neuroleptics and the pharmacological basis of schizophrenia.

This study was supported by the Wellcome Trust, MRC and the Research Funds of the Bethlem Royal and Maudsley Hospital and King's College Hospital. We thank Smith, Kline and French and Lundbeck for generous drug supplies.

a parallel experiment using inulin-activated serum, however, ^{125}I - β -endorphin was also recovered in the fractions (22–23S) known to contain the SC5b–9 complement complex³.

The identity of SC5b–9 with a β -endorphin-binding component in inulin-activated serum was confirmed by the demonstration that purified SC5b–9 exhibited β -endorphin-binding capacity in the radioreceptor assay (Fig. 1). ^{125}I - β -endorphin also co-sedimented with isolated SC5b–9 in sucrose density gradients (not shown). Further experiments showed that human ^{125}I - β -endorphin was almost completely displaced from its binding sites on the SC5b–9 complement complex by unlabelled human β -endorphin (Fig. 1).

Saturable binding sites for human β -endorphin were also found on the SC5b–8 and SC5b–7 complexes of human complement, but not on the isolated C3 or C9 components (L.S., S.B. and H.T., unpublished observations).

^{125}I - β -endorphin also bound to the isolated, cytolytic C5b–9(m) complement complex, whereas no binding was observed to hypotonically lysed sheep erythrocyte membranes (radio-receptor assay data are shown in Fig. 1). Consistent with this finding, total radioactivity was recovered from the top fractions of density gradients when β -endorphin binding to control mem-

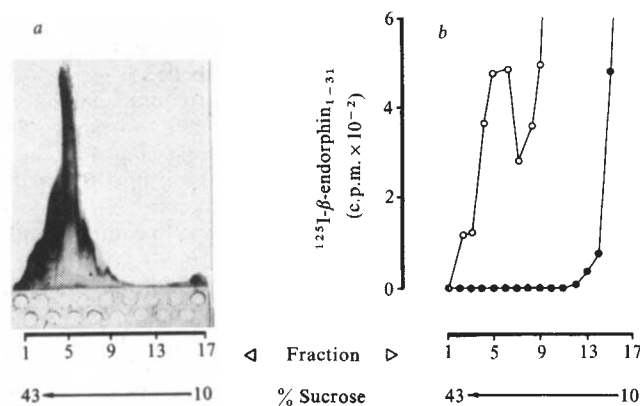


Fig. 2 Binding of human ^{125}I - β -endorphin_{1–31} to C5b–9(m) demonstrated by co-sedimentation in a sucrose density gradient. *a*, Fused rocket immunoelectrophoresis²² (using a specific anti-serum to identify C5b–9(m) complex) from fractions of density gradient centrifugation of a sample containing C5b–9(m) complex. *b*, Radioactivity determined in the fractions from a density gradient centrifugation of samples containing ^{125}I - β -endorphin incubated with either C5b–9(m) complex (○) or detergent-solubilized control erythrocyte membranes (●). No radioactivity sedimented in the control experiment. In the experiment performed with C5b–9(m) complex, however, the radioactivity peak coincided with the peak of immunoprecipitate; no protein was found in the upper fractions (9–13) of the gradient, indicating that the radioactivity in these fractions reflects ^{125}I - β -endorphin dissociated during centrifugation. For density gradient centrifugation 50 μl ^{125}I - β -endorphin solution (4 nM) were incubated at 22 °C for 20 min with either 250 μl inulin-activated serum or SC5b–9 or C5b–9(m) complex, or, as controls, with saline, EDTA-plasma or erythrocyte membranes. Samples were then centrifuged at 8 °C through 12.4 ml linear 10–43% (w/v) sucrose density gradients (containing 25 mM Tris buffer, 50 mM NaCl, 5 mM NaN₃, 2.4 mM Triton X-100, 2.5 mM DOC pH 8.2) at 35,000 r.p.m. in a SW41 Ti rotor (Beckman L2 65 B) for 16 h (ref. 3). Fractions were collected from the bottom of the tube and assayed for radioactivity.

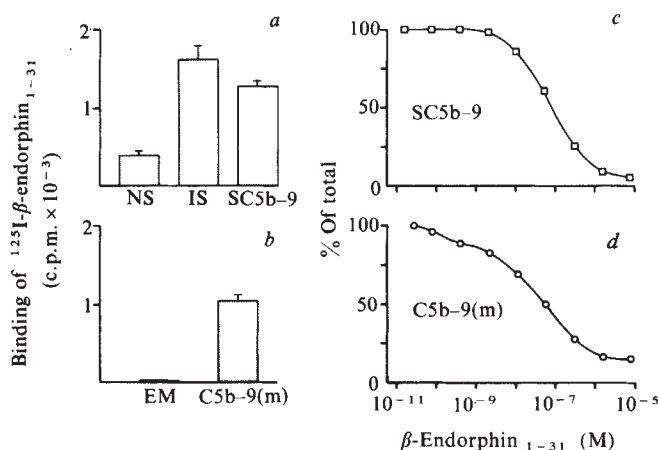


Fig. 1 *a, b*, Binding of human ^{125}I - β -endorphin_{1–31} (c.p.m.) to native human serum (NS); inulin-activated serum (IS); the cytolytically inactive, terminal complex of human complement (SC5b–9); control erythrocyte membranes (EM); and the terminal, cytolytic complex of human complement [C5b–9(m)]. Mean values $\pm$ s.d. from three determinations. *c, d*, Binding of human ^{125}I - β -endorphin_{1–31} to the terminal SC5b–9 or C5b–9(m) complexes of human complement in the presence of various concentrations of unlabelled human β -endorphin_{1–31} (binding expressed as per cent of total binding, specific plus nonspecific, as determined in the absence of β -endorphin_{1–31}; 100% corresponds to the binding values for SC5b–9 or C5b–9(m), respectively, shown to the left). Binding was determined by radioreceptor assay (a modified radioimmunoassay technique^{5,6}): 150 μl buffer (0.02 M sodium phosphate, 0.15 M NaCl, 0.01% bovine serum albumin (BSA), 0.1% gelatin, 0.01% thiomersol, pH 7.5) were incubated with 50 μl serum or 50 μl complement complex or membrane solution (adjusted to 300 or 30 μg protein per ml buffer for displacement or saturation binding experiments respectively) and with 20 μl buffer containing ^{125}I - β -endorphin and 0.1% Triton X-100. In addition, 20 μl buffer containing 0.1% Triton X-100 were added; in the competition experiments, the buffer also contained unlabelled inhibitor. Total concentration of ^{125}I - β -endorphin in radioreceptor assays was adjusted to $\sim 8,000$ c.p.m. per tube corresponding to 0.35 nM. The samples were incubated at 4 °C for 18 h. Then 300 μl of ice-cold charcoal suspension (1 g charcoal and 0.5 g BSA in 100 ml buffer) were added to adsorb free labelled or unlabelled ligands. After 10 min incubation at 4 °C the samples were centrifuged and the radioactivity in the supernatants was measured as an indication of the amount of ^{125}I - β -endorphin bound to complement complex.

branes was tested (Fig. 2); however, on centrifugation of the C5b–9(m) complex incubated with ^{125}I - β -endorphin, radioactivity was also found in the fractions (25–40S) which contained the complex⁷ (Fig. 2). Figure 1 shows displacement of ^{125}I - β -endorphin from the C5b–9(m) complement complex by unlabelled β -endorphin.

To test the reversibility of the binding reaction, we performed chase experiments. Samples containing C5b–9(m) were incubated with ^{125}I - β -endorphin for 18 h at 4 °C. Subsequent addition of unlabelled β -endorphin resulted in a decrease of ^{125}I - β -endorphin binding providing evidence for the reversibility of ^{125}I - β -endorphin binding to the C5b–9(m) complex. The shapes of the displacement curves (Fig. 1*d*; Fig. 4) and a Scatchard analysis of binding data (Fig. 3*b*) suggested negative cooperativity or several classes of binding sites. Assuming two classes of binding sites, we calculated two corresponding K_d values from the Scatchard plot (Fig. 3*b*) of about 80 and 3 nM respectively; the latter was in reasonable agreement with a K_d of about 1 nM estimated from the binding curve of a saturation experiment (Fig. 3*a*).

Classical opiate receptors are known to bind the N-terminal fragment of β -endorphin: β -endorphin is displaced from these binding sites by opioid peptides such as Met- and Leu-enkephalin, and by the opiate antagonist naloxone. None of these substances, in concentrations up to 10 μM , was able to inhibit ^{125}I - β -endorphin binding to the C5b–9(m) complex. Thus the N-terminal fragment of β -endorphin is obviously not essential for the binding of human β -endorphin to the human complement complex C5b–9(m).

To determine which fragment of human ^{125}I - β -endorphin might interact with the C5b-9(m) complex, competition experiments were performed using several unlabelled fragments of human β -endorphin and camel β -endorphin as inhibitors (Fig. 4). The amino acid sequences of camel and human β -endorphin differ only at positions 27 and 31 (see Fig. 4 legend). β -Endorphin $_{1-16}$, human β -endorphin $_{1-27}$ and camel β -endorphin $_{1-31}$ all failed entirely to compete with human ^{125}I - β -endorphin $_{1-31}$ in binding to the C5b-9(m) complex. In contrast, both human β -endorphin $_{1-31}$ and human β -lipotropin $_{1-91}$ which contains the amino acid sequence of human β -endorphin between positions 61 and 91, displaced ^{125}I - β -endorphin $_{1-31}$. Thus, the C-terminal fragment of human β -endorphin seemed to be critical for interaction between β -endorphin and C5b-9(m). Indeed, further experiments performed with the C-terminal fragment Lys-Lys-Gly-Glu of human β -endorphin showed that this tetrapeptide was capable of displacing ^{125}I - β -endorphin $_{1-31}$ from the C5b-9(m) complex (Fig. 4); however, the concentrations required were about 1,000-fold higher than those of human β -endorphin $_{1-31}$. All the data therefore support the existence of novel, non-opiate binding sites on the C5b-9(m) complex for the C-terminal fragment of β -endorphin. The relatively high concentrations of the C-terminal tetrapeptide required to displace ^{125}I - β -endorphin suggest that other regions of the peptide may also contribute to the binding.

The structural specificity of β -endorphin binding was further confirmed by the demonstration that ^{125}I - β -endorphin was not displaced from the C5b-9(m) complex by other peptides (10 μM) such as adrenocorticotropin $_{1-24}$, insulin, luteinizing hormone-releasing factor, dynorphin $_{1-13}$ or β -casomorphin $_{1-7}$. Human leukocyte interferon 8 and IgG 9 have been reported to interact with endorphin antisera; however, neither human fibroblast interferon nor IgG and its Fab or Fc fragments affected the binding of ^{125}I - β -endorphin to the C5b-9(m) complex.

The binding of human β -endorphin to the terminal C5b-9(m) complement complex thus exhibits several characteristics of a ligand-receptor interaction: reversibility, structural selectivity, saturability and high affinity; and it is tempting to assume that the process is biologically significant. Indeed recent results are consistent with an involvement of endorphins in the immune system: a transport system for benzomorphans in leukocytes has been reported 10 ; opioids apparently influence plasma levels of interferon 11 ; α -endorphin-immunoreactive material has been detected in plasma cells 12 ; and there is evidence for the

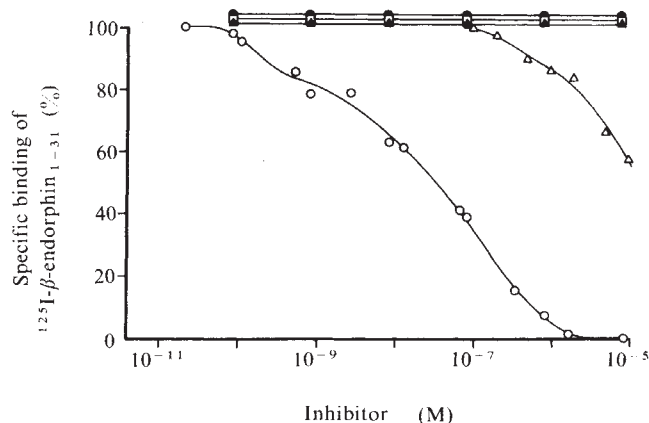


Fig. 4 Specific binding of human ^{125}I - β -endorphin $_{1-31}$ to C5b-9(m) in the presence of various concentrations of unlabelled binding inhibitors: $\square$, β -endorphin $_{1-16}$; $\blacktriangle$, human β -endorphin $_{1-27}$; $\bullet$, camel β -endorphin $_{1-31}$ (C-terminal amino acid sequence His-Lys-Lys-Gly-Gln); $\circ$, human β -endorphin $_{1-31}$ (C-terminal amino acid sequence Tyr-Lys-Lys-Gly-Glu); and $\triangle$, human β -endorphin $_{28-31}$ (amino acid sequence Lys-Lys-Gly-Glu). Mean values from three determinations (s.d. $\pm$ 4%). Specific binding represents bound ^{125}I - β -endorphin, that is displaced by 8.3 μM unlabelled human β -endorphin $_{1-31}$.

presence of opioid receptors on phagocytic leukocytes 13 and T lymphocytes 14,15 as well as for non-opiate β -endorphin receptors on cultured lymphocytes 16 .

β -Endorphin appears to be released from the pituitary into the blood under physical or emotional stress $^{17-19}$ and during parturition 20,21 . This might represent an adaptation of the immune system to stress situations.

We thank Margit Roth, Mechthild Vogel and Silvia Wilhelm for technical assistance, Professor Glossmann for valuable discussions, Dr Li for human β -lipotropin $_{1-91}$ samples and Dr Hofstetter for the gift of IgG and its Fab and Fc fragments. This study contains parts of the thesis of L. Schweigerer to be presented to the Fachbereich Humanmedizin in partial fulfilment of a doctorate of medical science. The work was supported by Deutsche Forschungsgemeinschaft (SP Neuroendokrinologie and grant We 340/10).

Received 28 September 1981; accepted 17 February 1982.

- Amir, S., Brown, Z. W. & Amit, Z. *Neurosci. Biobehav. Rev.* **4**, 77-86 (1980).
- Kolb, W. P. & Müller-Eberhard, H. J. *J. exp. Med.* **141**, 724-735 (1975).
- Bhakdi, S. & Roth, M. *J. Immun.* **127**, 576-602 (1981).
- Tranum-Jensen, J., Bhakdi, S., Bhakdi-Lehnen, B., Bjerrum, O. J. & Speth, V. *Scand. J. Immun.* **7**, 45-56 (1978).
- Höllt, V., Przewlocki, R. & Herz, A. *Naunyn-Schmiedeberg's Arch. exp. Path. Pharmac.* **303**, 171-174 (1978).
- Guillemin, R., Ling, N. & Vargo, T. *Biochem. biophys. Res. Commun.* **77**, 361-366 (1977).
- Bhakdi, S. & Tranum-Jensen, J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1818-1822 (1981).
- Blalock, J. E. & Smith, E. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5972-5974 (1980).
- Julliard, J. H., Shibasaki, T., Ling, N. & Guillemin, R. *Science* **208**, 183-185 (1980).
- Marks, M. J. & Medzihradsky, F. *Molec. Pharmac.* **10**, 837-848 (1974).
- Geber, W. F., Lefkowitz, S. S. & Hung, C. Y. *Pharmacology* **14**, 322-329 (1976).
- Grube, D. *Histochemistry* **69**, 157-160 (1980).
- Lopker, A., Abood, L. G., Hoss, W. & Lionetti, F. J. *Biochem. Pharmac.* **29**, 1361-1365 (1980).
- McDonough, R. J. *et al. J. Immun.* **125**, 2539-2543 (1980).
- Wybran, J., Appelboom, T., Famaey, J. P. & Govaerts, A. *J. Immun.* **123**, 1068-1070 (1979).
- Hazum, E., Chang, K. J. & Cuatrecasas, P. *Science* **205**, 1033-1035 (1979).
- Colt, E. W. D., Wardlaw, S. L. & Frantz, A. G. *Life Sci.* **28**, 1637-1640 (1981).
- Fraaioli, F. *et al. Experientia* **36**, 987-989 (1980).
- Teschemacher, H., Breidenbach, T., König, A., Luckhardt, M. & Davies-Osterkamp, S. in *Endogenous and Exogenous Opiate Agonists and Antagonists* (ed. Way, E. L.) 307-308 (Pergamon, New York, 1980).
- Akil, H., Watson, S. J., Barchas, J. D. & Li, C. H. *Life Sci.* **24**, 1659-1666 (1979).
- Csontos, K. *et al. Life Sci.* **25**, 835-844 (1979).
- Svendsen, P. J. *Scand. J. Immun. Suppl.* **1**, 69-70 (1973).
- Boeynaems, J. M. & Dumont, J. E. *Outlines of Receptor Theory* (Elsevier, Amsterdam, 1980).

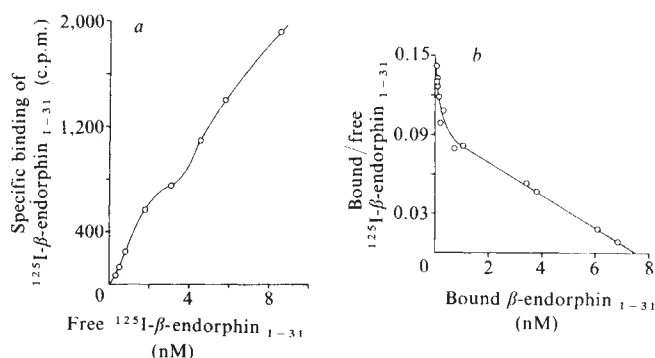


Fig. 3 *a*, Saturation experiment. Increasing concentrations of human ^{125}I - β -endorphin $_{1-31}$ were incubated with C5b-9(m) (6 μg protein per ml) in the presence or absence of 1.7 μM human β -endorphin $_{1-31}$. Free ^{125}I - β -endorphin was plotted against specifically bound (displaceable by 1.7 μM unlabelled β -endorphin) ^{125}I - β -endorphin. Mean values from four determinations (s.d. $\pm$ 4%). Biphasic curves were obtained in several experiments. *b*, Scatchard analysis of binding data obtained by displacement 23 of human ^{125}I - β -endorphin by unlabelled human β -endorphin (Fig. 4). Specifically bound/free ^{125}I - β -endorphin is plotted as a function of total (labelled plus unlabelled) ligand specifically bound.

Parental exposure to X rays and chemicals induces heritable tumours and anomalies in mice

Taisei Nomura

Institute for Cancer Research, 1st Department of Surgery, and Department of Fundamental Radiology, Osaka University Medical School, Fukushima-ku, Osaka 553, Japan

Although radiation and various chemicals are known to produce mutations, there is a paucity of information on the production of germ-line mutations leading to heritable tumours and anomalies. I now report a significant and large increase in such tumours and anomalies following treatment of ICR parent mice with X rays and urethane. About 90% of the induced tumours were in the lung, and were inherited as if they were dominant mutations with about 40% penetrance.

X-ray irradiation was chosen because of the substantial previous experimental work on radiation mutagenesis¹⁻⁶. Urethane is known to be mutagenic, carcinogenic and teratogenic in mice and has also been used widely as a human drug⁷. Additional experiments, with similar results, were done with 4-nitroquinoline 1-oxide (4-NQO). The present report summarizes the data from large-scale experiments involving 12,905 live-born mice and 9,645 fetuses from 2,904 parents studied during 1967-81 (the spontaneous and induced rates of tumours and anomalies have not changed noticeably in this period). The experimental procedures and detailed results will be published elsewhere.

The first comparison between radiation and urethane was in the production of dominant lethals. As shown in Fig. 1, there was a significant, almost linear increase in dominant lethals following irradiation of spermatids and spermatozoa. Similarly, there was an increase of dominant lethals following treatment of mature oocytes, with no significant fractionation effect (although the data, taken at face value, suggest a slight reduction with fractionation). On the other hand, treatment of spermatogonia produced no significant increase, although the numbers are too small for a small change to be detected. These results are in accord with previous studies⁸⁻¹¹. In contrast, examination of 5,830 embryos revealed no significant increase of dominant lethals from urethane treatment (1.5, 2.0 and 2.25 mg per g) at any stage. Nevertheless, as will be shown below, urethane is very effective in producing tumours and anomalies.

As shown in Table 1, there is a large and significant increase in anomalies in the progeny of parents treated with radiation or urethane. Generally, a higher rate of anomalies is detected prenatally than after birth, because many of the anomalies (for example, cleft palate, exencephalus, gastroschisis and buphthalmus) are lethal shortly after birth. Only open eyelid, the pre-dominant type among non-lethal anomalies, was shown to be transmissible (see Table 1 legend).

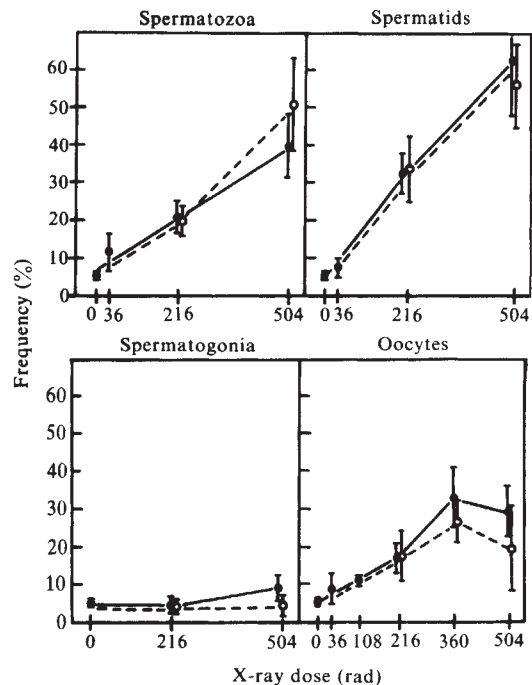


Fig. 1 Per cent of dominant lethals in the progeny of mice irradiated at the stages and dosages shown. ●, Acute doses; ○, fractionated doses, given in 36 rad amounts at 2-h intervals. Dominant lethals were measured by the mean number of early deaths per implant. Vertical bars indicate the 90% confidence limits of the mean computed from the *t* distribution. For spermatozoa treatment males were mated 1-7 days after irradiation, for spermatid 15-21 days, and for spermatogonia 64-80 days; for mature oocytes at the late follicular stage, females were irradiated 1-14 days before ovulation.

mus) are lethal shortly after birth. Only open eyelid, the pre-dominant type among non-lethal anomalies, was shown to be transmissible (see Table 1 legend).

Figure 2 gives the dose-response data for radiation and the effects of dose-fractionation. The data are too limited to permit any conclusions about the shape of the dose-response relationship for male treatment, although there is a clear increase with dose in females. As with dominant lethals, the treatment of spermatogonia was less effective than post-meiotic treatment. There is again a suggestion of a fractionation effect for oocyte treatment.

Both radiation and urethane significantly increase the number of tumours in the progeny, regardless of which sex is treated (Table 1). The dose-response data for radiation (Fig. 3) show a clear increase with dose for treatment of post-meiotic stages in males and less clear results for spermatogonial treatment. Oocytes at late follicular stages were resistant to 36-108 rad of X rays, but very sensitive to higher doses. Both the

Table 1 Incidence of tumours and anomalies in the offspring of mice exposed to X rays or urethane

Agent	Sex	Treatment to parent Dose (rad or mg per g)	Anomalies detected in 19-day-old fetuses		Anomalies detected in 7-day-old offspring		Tumours in offspring 8 months after birth	
			Incidence	(%)	Incidence	(%)	Incidence	(%)
X rays	M	36-504	48/2,201	(2.2*)	16/2,396	(0.7)	153/1,529	(10.0†)
X rays	F	36-504	25/942	(2.7*)	29/1,750	(1.7†)	101/1,155	(8.7‡)
Urethane	M	2.0	10/477	(2.1†)				
Urethane	M	1.5	65/2,923	(2.2*)	16/1,637	(1.0)	136/1,254	(10.9†)
Urethane	F	1.5	52/1,262	(4.1*)	22/1,087	(2.0*)	139/963	(14.4*)
Urethane	F	1.0			12/885	(1.4†)	115/772	(14.9*)
Untreated		0.0	4/1,026	(0.4)	1/809	(0.1)	29/548	(5.3)

X-ray irradiation was by Toshiba KC-18-2A, 180 KVP at a rate of 72 rad min⁻¹. Data are totals for various dosages, fractionations and age of exposure. Urethane was given as a single subcutaneous dose. Pathology was detected both by gross examination and histologically. The induced anomalies were: cleft palate, kinky and/or short tail, dwarf, open eyelid, exencephalus, hydrocephalus, gastroschisis, polydactyly, syndactyly, gigantic toe, buphthalmus, hydatid mole, atresia hymenalis, mislobulation of lung or liver, and hemiplegia. Among non-lethal anomalies, open eyelid, the predominant type (41%), was transmissible to F₂, but transmission of tail anomaly was inconsistent. Other non-lethal anomalies are sterile or too rare to examine exact heritability. Types of induced tumours are given in the text. Results were significantly different from the control value at **P* < 0.001, †*P* < 0.01 and ‡*P* < 0.05 by χ^2 .

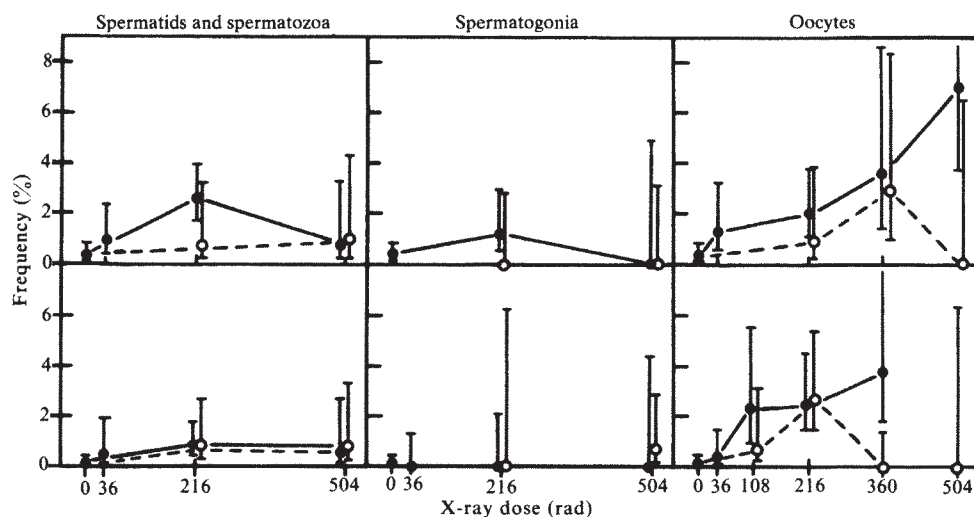
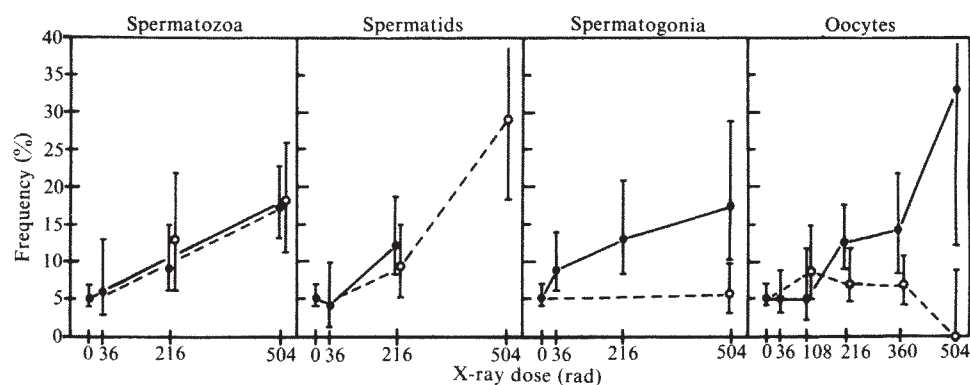


Fig. 2 Frequency of prenatal (above) and postnatal anomalies (below) in the progeny of parents irradiated in a single dose (●) and fractionated (○). Frequency of the anomaly was given by per cent of anomalous progeny in survivors. Vertical bars indicate 90% binomial confidence limits. For treatments, see Fig. 1 legend. Although data are not shown, spermatocytes at 22–42 days before mating were twice as sensitive to X rays as were post-meiotic sperm. Further study on dose-response relationship remains with these stages.

Fig. 3 Frequency of tumours in the progeny of mice irradiated at various stages and doses. ●, Acute doses; ○, fractionated doses. Tumour frequency was given by per cent of tumour-bearing progeny in survivors. Vertical bars indicate 90% binomial confidence limits. For treatments, see Fig. 1 legend.



spermatogonial and oocyte treatment effects of dose-fractionation suggest some repair at high doses. Eighty-seven per cent of the induced tumours were in the lung (papillary adenomas), the remainder being of various types—ovarian tumour, lymphocytic leukaemia, stomach tumour, lipoma, granulosa cell tumour, thyroid tumour, liver haemangioma and hepatoma. The spectrum of tumour incidence in the offspring of treated animals is essentially the same as in controls (90% in the lung).

To confirm that the induced tumours are heritable, the F_1 progeny of treated parents were mated and their progeny examined. In addition to those tumours induced by urethane and radiation, some induced by 4-NQO were included. The mice were mated as young adults and the presence or absence of tumours was determined at 8 months of age. The progeny were then classified as to the retrospectively determined type of the parent. The data are shown in Table 2. In each case the original treated mouse was a male, as a precaution against the possible transmission of the chemical to the progeny. The results were confirmed by continuing the urethane-treated group into the F_3 generation. The F_3 results were: $+ \times -$, 5/22; $- \times -$,

5/125 (see Table 2 legend for explanation of symbols). Combining the data, the proportion of affected progeny from $+ \times -$ matings was 21.4%. The pattern of inheritance is that of a dominant with about 40% penetrance. Reduced penetrance was also found for dominant skeletal mutations⁶.

In addition, four untreated females that later developed spontaneous lung tumours produced 49 progeny. Of these, only two (4.1%) developed tumours, a value not different from the progeny of normal mice in this strain. The histological pattern was the same in the two groups. One possible explanation for the different heritability is that the spontaneous tumours were caused by somatic rather than germinal mutation. In view of the quantitative doubts about spontaneous rates, doubling dose calculations are not presented.

Although tumours were produced by both radiation and urethane, urethane failed to produce a significant increase of dominant lethals, suggesting that gross chromosomal rearrangements are not produced by urethane. To test this further, progeny of treated parents were examined for translocations detected cytologically at meiosis. Of 64 progeny from parents treated with 504 rad of X rays at post-meiotic stages, 10 had characteristic translocation configurations. Two others had small testes with no meioses, probably also caused by gross chromosomal rearrangements. One of these two had a lung tumour. Among the 54 progeny mice showing no evidence of translocations, eight had lung tumours. In contrast, no evidence of translocations was found in 50 progeny of parents treated with urethane. This difference between X rays and urethane was confirmed by genetic experiments with *Drosophila melanogaster*⁷. Urethane produced no chromosome rearrangements or dominant lethals, but was a potent mutagen. At the highest doses, the number of recessive X-linked lethals was equivalent to those produced by 2,000 rad of γ rays.

Table 2 Lung tumours in F_2 progeny of treated parents

Phenotypes of F_1 parents	Parental treatment		
	X rays	4-NQO	Urethane
$+ \times -$	15/76 (19.7*)	7/28 (25.0†)	4/19 (21.2)
$- \times -$	4/69 (5.8)	5/103 (4.9)	9/113 (8.2)

Male parents were treated with X rays (504 rad), 4-NQO (10 μ g per g) or urethane (1.5 mg per g) at 8–14 days before mating. + Means that the F_1 mouse later developed lung tumour; - means that it had not developed a tumour by 8 months. There was significant difference between $+ \times -$ and $- \times -$ mating at * $P < 0.05$ and † $P < 0.01$.

The fact that urethane produces tumours but not translocations or dominant lethals and the observation that tumours occur no more frequently in those mice with X-ray-induced translocations (1/12) than in those without translocations (8/54) argues that the mutations producing tumours are not related to gross chromosome changes. There could, of course, be cytological changes (small deletions, for example) not detected as translocations. In any case, the tumours behave as dominant mendelian factors with a penetrance of about 40%. Reduced penetrance is not surprising in this kind of trait.

It is striking that tumour mutations occur about 10 times as frequently as dominant skeletal mutations^{4,5}. The spontaneous tumour incidence is also much higher (~5%) than skeletal mutations. It is possible that many more gene loci are involved¹²; alternatively, specific loci may be especially prone to lung tumour-causing mutations in my ICR strain.

The most important reservation regarding these experiments

is that the results may be a special property of the mouse strain used. However, other kinds of tumours (ovarian tumours and lymphocytic leukaemias), although much less frequent than lung tumours, increased in the same ratio. It is important that similar experiments should be done with other mouse strains to test the generality of the results. Confirmation of the findings in other strains would mean that heritable cancer induced by radiation and chemical mutagens becomes an important factor for study in human populations such as those of Hiroshima and Nagasaki. Note that in traditional experiments designed to test mutability in mice the progeny are not kept long enough for tumours to develop.

I thank Dr J. F. Crow for advice and help in summarizing this work, the late Dr J. Kamahora, and Drs Y. Sakamoto, S. Kondo and E. Okamoto for their encouragement, and Drs A. G. Searle and P. C. Hoppe for their critical comments. The work was supported by JMESC grant 56480373.

Received 10 November 1981; accepted 26 January 1982.

1. Russell, W. L. in *Repair from Genetic Radiation* (ed. Sobels, F. H.) 205-217 (Pergamon, Oxford, 1963); in *Proc. XI Int. Congr. Genetics*, 257-264 (Pergamon, Oxford, 1964).
2. Russell, W. L., Russell, L. B. & Kelly, E. M. *Science* **128**, 1546-1550 (1958); *Int. J. Rad. Biol. (Suppl.)*, 311-320 (1960).
3. Russell, W. L., Russell, L. B. & Oakberg, E. F. in *Radiation Biology and Medicine* (ed. Claus, W. D.) 189-205 (Addison-Wesley, Massachusetts, 1958).
4. Ehling, U. H. *Genetics* **54**, 1381-1389 (1966).

5. Bartsch-Sandhoff, M. *Humangenetik* **25**, 93-100 (1974).
6. Selby, P. B. & Selby, P. R. *Mutat. Res.* **43**, 357-375 (1977); **50**, 341-351 (1978); **51**, 199-236 (1978).
7. Nomura, T. *Cancer Res.* **34**, 2217-2231 (1974); **35**, 264-266, 2895-2899 (1975); **39**, 4224-4227 (1979); *Br. J. Cancer* **33**, 521-532 (1976).
8. Searle, A. G. & Beechey, C. V. *Mutat. Res.* **22**, 63-72 (1974).
9. Russell, L. B. & Major, M. *Genetics* **38**, 687-688 (1953).
10. Russell, W. L., Russell, L. B. & Kimball, A. W. *Am. Nat.* **88**, 269-286 (1954).
11. Bateman, A. J. *Heredity* **12**, 213-232 (1958); *Mutat. Res.* **4**, 710-712 (1967).
12. Flavell, A. *Nature* **289**, 10-11 (1981).

A host gene controlling early anaemia or polycythaemia induced by Friend erythroleukaemia virus

Tsunefumi Shibuya & Tak W. Mak

The Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, 500 Sherbourne Street, Toronto, Ontario, Canada M4X 1K9

Friend erythroleukaemia virus (FLV) is a type-C murine retrovirus that induces a rapid multi-stage erythroleukaemia in susceptible mice¹⁻⁵. An early phase (1-3 weeks) is characterized by splenomegaly and rapid changes in erythropoiesis⁵⁻⁸, while a later stage (1-3 months) is associated with the development of clones of highly tumorigenic cells^{2-5,9}, which can be established into Friend erythroleukaemic cell lines. Two distinct isolates of FLVs, both of which can induce the early and late stages of the diseases, have been described. The original isolate (FV-A)¹ induces a rapid severe anaemia¹, while a later isolate (FV-P)¹⁰, derived from stocks of FV-A, induces a rapid polycythaemia⁷. In addition to the contrasting changes in haematocrit, the diseases induced by FV-A and FV-P can also be distinguished. For example, while the erythropoiesis in FV-A infected mice appears to remain under the control of the hormone erythropoietin^{8,11,12}, erythropoiesis in FV-P infected mice is erythropoietin-independent^{8,11-13}. Furthermore, Friend erythroleukaemic cell lines derived as a consequence of FV-A and FV-P infection also display different cellular properties^{5,9,14}. Here we demonstrate that in addition to the viral contributions, host genes play an important part in determining the nature of the disease induced by FV-A or FV-P. A new host gene which determines the induction of early anaemia or polycythaemia by FLV is described.

The polycythaemia strain of FV-P has been used extensively in studies of host genetic control as well as its interaction with the haematopoietic progenitor cells^{15,16}. Most of such studies used mice strains DBA/2 and BALB/c. We have now extended these studies to other mice strains, notably CBA. Figure 1a

illustrates the change in haematocrit in DBA/2 mice following the infection by FV-A and FV-P. The data show that for FV-A, the haematocrit of the mice dropped from a normal value of 50% to ~35% by 2-3 weeks, while that of FV-P infected mice increased to over 70% during the same period. However, as also shown in Fig. 1a, following the infection of FV-P into another strain of mouse CBA, instead of increasing, the haematocrit fell to ~35%. A similar reduction of the proportion of red cells in the circulating blood was also detected following infection with FV-A. This implies that not only can a contrasting response in haematocrit values be stimulated by FV-A and FV-P, but that different haematocrit values can be induced after infection of FV-P into two different strains of mice. These rapid changes in haematocrit appear to be the effect of the replication-defective spleen focus-forming virus components, SFFV_A and SFFV_P (ref. 8) of FV-A and FV-P, respectively. The helper virus, F-MuLV, did not induce such changes in either DBA/2 or CBA mice (Fig. 1a).

It is possible that CBA mice are resistant to polycythaemia induction by FV-P and that the early anaemia observed in these mice after injection of FV-P is due to a small quantity of contaminating SFFV_A in the FV-P stocks. To rule out this possibility, we have prepared a FV-P preparation by transfection of molecularly cloned polycythaemia strain of spleen focus forming virus (SFFV_P) DNA¹⁷ to NIH/3T3 cells and rescued these transfected cells with cloned F-MuLV¹⁷. Data in Fig. 1b indicate that this virus preparation, which contained only virus from molecularly cloned SFFV_P and F-MuLV, also induced an early polycythaemia (70% haematocrit) in DBA/2 and a rapid anaemia (30% haematocrit) in CBA mice. Another possibility is that inoculation of FV-P into CBA mice resulted in certain host-induced changes in the viral genomes, or that endogenous FV-A was induced, thus causing anaemia in these mice. To rule this out, a 10% spleen extract was obtained from the CBA mice 2 weeks after infection with FV-P and injected into five DBA/2 and five CBA mice. Results indicate that, similar to the original stock of FV-P, the virus in these spleen extracts induced a severe polycythaemia (average haematocrit 69 ± 2.5%) in DBA/2 mice and a rapid anaemia (haematocrit 40 ± 4%) in CBA mice 2 weeks after inoculation.

We next examined several inbred mouse strains and their changes in haematocrit after infection by FV-P. Mouse strains DBA/2, CBA, C3H, AKR and BALB/c were injected with an equal dose of FV-P and their haematocrits measured at the

time of infection (0 week) and 2 weeks later. Results in Table 1 indicate that, like that of DBA/2 mice, the AKR mice and BALB/c mice strains developed rapid polycythaemia. The C3H strain, however, responded like CBA mice in developing an anaemia. These contrasting results demonstrate that host determinants have an important role in determining these early erythropoietic changes.

To examine if such contrasting effects in the changes of haematocrit as a result of infection by FV-P is the result of a single locus gene, (CBA×DBA/2) F₁ hybrid mice and mice from backcrosses of DBA/2×(CBA×DBA/2) F₁ and CBA×(CBA×DBA/2) F₁ were developed and their haematocrit changes 2 weeks after injection with FV-P were measured. The data in Fig. 2 indicate that, similar to the findings described above, all the DBA/2 mice injected with FV-P developed a rapid polycythaemia with a haematocrit value of about 65–70%, while the haematocrits of all the CBA mice dropped to 35–40%. The (CBA×DBA/2) F₁ hybrids, however, developed a slight polycythaemia, with haematocrit values between those of CBA and DBA/2 mice, 55–60%. These data suggest that if the early anaemia or polycythaemia is controlled by a single gene, the gene has a co-dominant effect. The haematocrits of mice from CBA×(CBA×DBA/2) F₁ appear to segregate into two groups, one group with haematocrits of 35–40% and another with values of 55–60%, while the haematocrits of the DBA/2×(CBA×DBA/2) F₁ mice seem to have segregated into two groups, one at 55–60% and one at 65–70%. These results indicate that the haematocrit values in the hybrid mice segregated as single genotypes, suggesting that the induction of early

Table 1 Early changes in haematocrit in different mouse strains following infection by FV-P

Mice	Haematocrit	
	0 week	2 weeks
DBA/2	51.0±1.0	66.5±0.9
BALB/c	52.0±1.5	57.3±3.1
AKR	49.3±0.9	71.0±2.4
CBA	49.5±0.5	36.8±3.7
C3H	49.8±0.5	40.0±1.9

2×10³ FFU of FV-P were injected into DBA/2, BALB/c, AKR, CBA or C3H mice and the haematocrit values were determined at the time of infection and 2 weeks later (see Fig. 1). All mice used were female, 7–10 weeks old at the time of infection. At least four mice were used in each determination. Error represents 1 s.d.

polycythaemia or anaemia in DBA/2 and CBA mice, respectively, is controlled by a single co-dominant locus.

The contributions of host hereditary factors in determining the susceptibility to FLV infection is well established^{14–16}. Several host genes—*Fv-1* (ref. 15), *Fv-2* (refs 16, 18, 19), *W* (refs 8, 20), *Steel* (refs 8, 16, 21, 22) and *H-2* (refs 16, 23)—are known to control the susceptibility of mice to erythroleukaemia induction by this virus. The present gene, however, seems to affect the type of erythropoietic modulations by FLV rather than the susceptibility.

This finding, taken together with the previous reports that the types of modification of the erythropoiesis patterns depend on the strain of FLV used^{8,11,12}, indicate that the modulation of the haematopoietic organization by FLV is the result of a complex array of interactions between host genes and the viral genomes. These early experiments showed that, following infection by FV-A and FV-P, certain pathological differences could be detected. For example, FV-P induces a rapid polycythaemia associated with erythropoietin-independent erythropoiesis^{8,11–13}, while FV-A induces a rapid anaemia associated with erythropoietin-dependent erythropoiesis^{8,11,12}. The reason for these contrasting changes is not known. Perhaps the two viruses infect different target cells in the haematopoietic hierarchy, leading to the modulation and appearance of cells with different properties. This hypothesis is also supported by the nature of the tumorigenic colonies emerging in the late stages of the disease^{5,9}. These colonies, which are probably the result of progressions from the infected cells in the early phase of the disease, when derived from FV-P- or FV-A-infected mice, differ in their kinetics of appearance⁵, their self-renewing capacities⁹, their incidences of chromosomal aberrations⁹ and their responsiveness to induction by erythropoietin^{9,14}.

The present data indicate that in addition to the contributions of host genes in determining the susceptibility to infection by FLV, the outcome of the virus-induced erythropoietic modulations is under genetic control. We propose that this locus in DBA/2 and CBA mice, which controls the state of erythropoiesis early after infection by FV-P, be designated *Fv-5*. Thus, CBA mice which develop early anaemia after infection with FV-P have a genotype of *Fv-5^{aa}* and DBA/2 mice which develop polycythaemia, *Fv-5^{pp}*. The mode of action of this host gene is unknown. Although it is possible that the viruses infect different target cells in CBA and DBA/2 mice, leading to the infection or transformation of non-erythroid cells or of erythroid cells with different differentiation potentials, preliminary results do not support this. Similar to DBA/2 mice infected with FV-P, the CBA spleen cells are predominantly erythroblasts. Furthermore, erythropoiesis of the FV-P-infected CBA mice and the formation of erythroid colony-forming units were also independent of any added erythropoietin (T.S., Y. Niho and T.W.M., unpublished data). Alternatively, this gene may affect the severity of haemolysis of the FLV-transformed erythroid cells, or control the rate of proliferation and differentiation of late erythroid cells. Finally, it may function by controlling the transport of the transformed erythroid cells from the bone

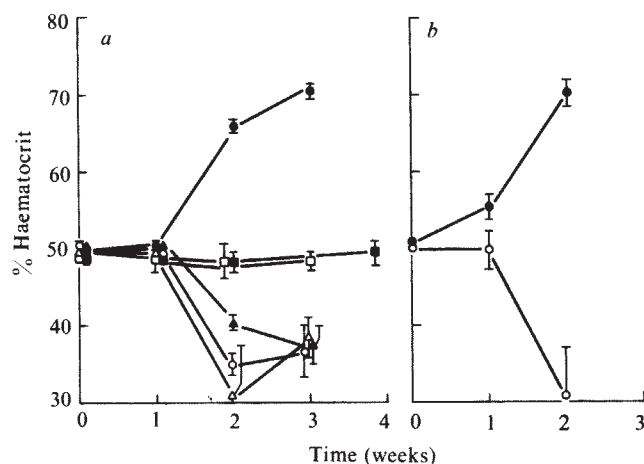


Fig. 1 *a*, Kinetics of changes in haematocrit values in DBA/2 and CBA mice infected with FV-P, FV-A or F-MuLV. Clonal isolates of FV-P, FV-A or F-MuLV used were described elsewhere⁸. To examine the kinetics of changes in haematocrit values, 2×10³ focus-forming units (FFU) of FV-P, 10³ FFU of FV-A, or 2.5×10⁶ XC plaque units of F-MuLV were injected through the tail veins of female DBA/2 mice or CBA female mice (Jackson Laboratories). At weekly intervals, the haematocrit values of the mice were determined by bleeding through the tail veins of the mice with haematocrit tubes (Dade Division, American Hospital Supply, Miami). At least four mice were used for each determination. All mice were 7–10 weeks old at the time of infection. Symbols: ●, FV-P into DBA/2; ○, CBA; ▲, FV-A into DBA/2 and CBA; △, ■, □ F-MuLV into DBA/2 or CBA. Bar represents 1 s.e. *b*, Changes of haematocrit values in DBA/2 and CBA mice after infection of viruses from molecularly cloned SFFV_p DNA and F-MuLV. Molecularly cloned polycythaemia strain of spleen focus-forming virus (SFFV_p) DNA were transfected to NIH/3T3 cells¹⁷. The transfected cells were rescued with cloned F-MuLV^{8,17}, the virus in the supernatant was injected into either DBA/2 or CBA mice and their haematocrit values were determined as described for *a*. Symbols: ●, DBA/2; ○, CBA mice. At least four mice were used in each determination. Bar represents 1 s.e.

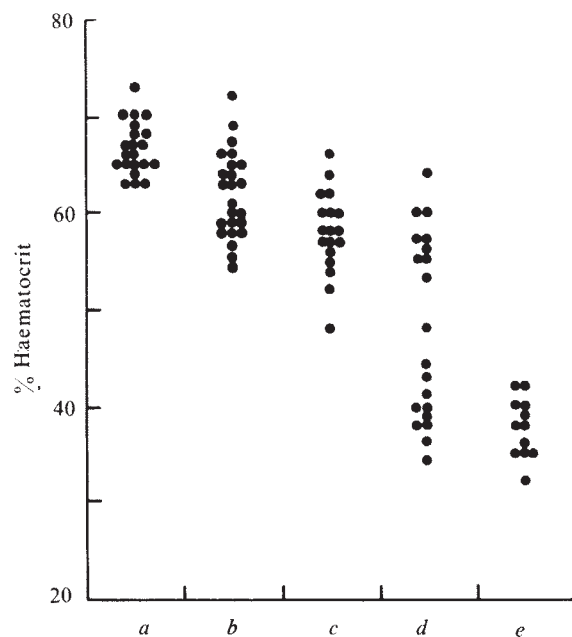


Fig. 2 Haematocrit values of: a, DBA/2; b, (CBA × DBA/2) F_1 ; c, (CBA × DBA/2) F_1 ; d, CBA × (CBA × DBA/2) F_1 ; e, CBA mice 2 weeks after infection by FV-P. 2×10^3 FFU of FV-P were injected into each mouse with background of DBA/2, CBA, (CBA × DBA/2) F_1 , DBA/2 × (CBA × DBA/2) F_1 and CBA × (CBA × DBA/2) F_1 , and their haematocrit values were determined 2 weeks after infection. (CBA × DBA/2) F_1 , DBA × (CBA × DBA/2) F_1 and CBA × (CBA × DBA/2) F_1 mice were developed in the animal colony of the Ontario Cancer Institute.

marrow or spleens of these mice, thus controlling the proportion of red cells in the circulation. Comparison of the properties of the erythroid progenitor cells, the rate of erythropoiesis and the nature of tumorigenic Friend erythroleukaemic cell lines derived after infection of mice of genotypes *Fv-5^{aa}*, *Fv-5^{ap}* and *Fv-5^{pp}* will help to elucidate the process of leukaemic transformation by FLV and cellular heterogeneity of leukaemic clones.

This work was supported by the MRC of Canada.

Note added in proof: Recent results from this laboratory suggest that *Fv-5* controls the rate of proliferation of a late (post CFU-E) erythroid precursor cell(s).

Received 12 August 1981; accepted 16 February 1982.

1. Friend, C. J. *exp. Med.* **105**, 307-318 (1957).
2. Friend, C. & Haddad, J. R. *J. natn. Cancer Inst.* **25**, 1279-1289 (1960).
3. Ostertag, W., Melderis, H., Steinheider, G., Kluge, N. & Dube, S. *Nature new Biol.* **239**, 231-234 (1972).
4. Mager, D., Mak, T. W. & Bernstein, A. *Nature* **288**, 592-594 (1980).
5. Mager, D. L., Mak, T. W. & Bernstein, A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1703-1707 (1981).
6. Tambourin, P., Wendling, F., Moreau-Gachelin, F., Charon, M. & Bucan-Varet, P. in *In Vivo and In Vitro Erythropoiesis: The Friend System* (ed. Rossi, G. B.) 127-138 (Elsevier, Amsterdam, 1980).
7. Horoszewicz, J. S., Leong, S. S. & Carter, W. A. *J. natn. Cancer Inst.* **54**, 265-267 (1975).
8. MacDonald, M. E. *et al. J. exp. Med.* **151**, 1477-1492 (1980).
9. Mager, D. L., MacDonald, M. E., Robson, I. B., Mak, T. W. & Bernstein, A. *J. molec. Cell Biol.* **1**, 721-730 (1981).
10. Mirand, E. A., Prentice, T. C. & Hoffman, J. G. *Proc. Soc. exp. Biol. Med.* **106**, 423-426 (1961).
11. Fagg, B., Vehmeyer, K., Ostertag, W., Jasmin, C. & Klein, B. in *In Vivo and In Vitro Erythropoiesis: The Friend System* (ed. Rossi, G. B.) 163-172 (Elsevier, Amsterdam, 1980).
12. Peschle, C. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 2054-2058 (1980).
13. Liao, S. K. & Axelrad, A. A. *Int. J. Cancer* **15**, 467-482 (1975).
14. Kluge, N., Gaedicke, G., Steinheidev, G., Duke, S. & Ostertag, W. *Expl Cell Res.* **88**, 257-262 (1974).
15. Steeves, R. & Lilly, R. A. *Rev. Genet.* **11**, 277-296 (1977).
16. Mak, T. W., Gamble, C. L., MacDonald, M. E. & Bernstein, A. *Cold Spring Harb. Symp. quant. Biol.* **44**, 893-898 (1980).
17. Yamamoto, Y. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 6893-6897 (1981).
18. Otake, T. *Int. J. Cancer* **16**, 18-23 (1970).
19. Lilly, F. J. *natn. Cancer Inst.* **45**, 163-169 (1970).
20. Steeves, R. A., Bennett, M., Mirand, E. A. & Cudkowicz, G. *Nature* **218**, 372-374 (1968).
21. Bennett, M., Steeves, R. A., Cudkowicz, G., Mirand, E. A. & Russell, L. B. *Science* **162**, 564-565 (1968).
22. McCool, D., Mak, T. W. & Bernstein, A. *J. exp. Med.* **149**, 837-846 (1979).
23. Lilly, F. J. *exp. Med.* **127**, 465-473 (1968).

Unequal crossing-over accounts for the organization of *Drosophila virilis* rDNA insertions and the integrity of flanking 28S gene

Peter M. M. Rae

Department of Biology, Yale University, New Haven, Connecticut 06511, USA

The majority of ribosomal DNA (rDNA) repeats in *Drosophila melanogaster* and *Drosophila virilis* have an insertion in the 28S rRNA coding region, and such repeats are essentially inactive in transcription¹. There are two unrelated types of insertion in *D. melanogaster* rDNA, one of which (type 1) has some homology with insertions in *D. virilis*², and is located in precisely the same position as that in the *D. virilis* rDNA³⁻⁶. *D. melanogaster* type 1 rDNA insertions fall into three size classes of 5, 1, 0.5 kilobases (kb), and sequences of shorter insertions comprise the downstream end of longer ones^{4,5}. Indeed, the points at which the sequences of longer type 1 insertions depart from shorter ones are oligonucleotides that are closely related to a 14-base pair (bp) rRNA coding sequence directly repeated at the flanks of *D. virilis*³ and some *D. melanogaster*⁴ rDNA insertions. It has been suggested that shorter *D. melanogaster* type 1 insertions derived from longer ones by unequal crossing-over that involved these homologies⁵, and I show here that long *D. virilis* rDNA insertions may have arisen from shorter ones in the same way. Also, I suggest that such unequal crossing-over can generate intact rDNA repeats from those with insertions.

The major *D. virilis* rDNA insertion has been described as being ~10 kb long^{2,3,7}, although it had been noted that it appears to be a tandem duplication². Blot hybridizations using uncloned *D. virilis* rDNA digested with restriction enzyme combinations not previously used have now shown that 5-kb insertions are at least as frequent as 10-kb insertions, and that a *Hind*III cleavage site was misinterpreted in a previous study². Restriction site mapping and the sequencing of homologous regions of a 10-kb insertion verified that it is a tandem duplication. This fact, together with sequence data on the middle of a 10-kb insertion, strongly suggest that it arose from 5-kb insertions by unequal crossing-over, with the flanking 14-bp repeats of coding sequence serving as points of homology. Longer segments containing the insertion have been found from blots to have lengths that are also multiples of ~5 kb. If these are from rDNA, they presumably arose from unequal crossing-over between rDNA repeats having insertions of various lengths. However, they may represent non-rDNA copies of the insertion^{1,2,6}.

The *D. virilis* rDNA clones pDvL1-2 and pDvS7-3 used in this study have been described elsewhere^{2,3,8}; the map of an rDNA repeat that contains a 10-kb insertion is shown in Fig. 1, together with the lengths of rDNA that were sequenced for this study.

Figure 2 (lane 2) shows the hybridization of DvS7-3, a *Bam*HI segment of *D. virilis* rDNA that contains only insertion sequences (Fig. 1), to *D. virilis* nuclear DNA digested with *Sma*I + *Hind*III and electrophoresed in a 0.4% agarose gel. There are no *Sma*I or *Hind*III sites in the known *D. virilis* rDNA insertions, but there is a *Sma*I site in the 28S' coding region ~2.2 kb upstream of the insertion (Fig. 1), and a *Hind*III site 280 bp downstream of the insertion in the 28S'' coding region³. The greatest amount of hybridization occurred to a band at 7.5 kb, and this length is attributable to a 5-kb insertion plus flanking coding sequences to the *Sma*I and *Hind*III sites. The next most intense band is at ~12.8 kb, and is considered to derive from rDNA repeats having 10-kb insertions. Other bands, indicated

Fig. 1 Maps of the cloned *D. virilis* rDNA segments DvL1-2 and DvS7-3 that together represent an rDNA repeat having a 10-kb insertion in the 28S rRNA coding region². DvL1-2 itself may be a segment derived from two adjacent repeats with 5-kb insertions, or from adjacent repeats with insertions of different size, but the contiguity in rDNA of *Bam*HI segments comparable with DvL1-1 and DvS7-3 has been demonstrated². The distribution of restriction enzyme cleavage sites in the insertion shown indicates that it is a tandem duplication, and a comparison of DNA sequences around the *Bam*HI, *Eco*RI and *Sal*I sites in DvL1-2 and DvS7-3 has verified this. The horizontal arrows beneath the expanded portions of the map indicate DNA that was sequenced specifically for this study. Sequencing was done according to Maxam and Gilbert¹³. DNA segments were either 5' end-labelled with [γ -³²P]ATP and polynucleotide kinase¹³, 3' end-labelled with [α -³²P]dNTP and the Klenow fragment of DNA polymerase¹⁴, or 3' end-labelled with [α -³²P]3'dATP and terminal transferase¹⁵.

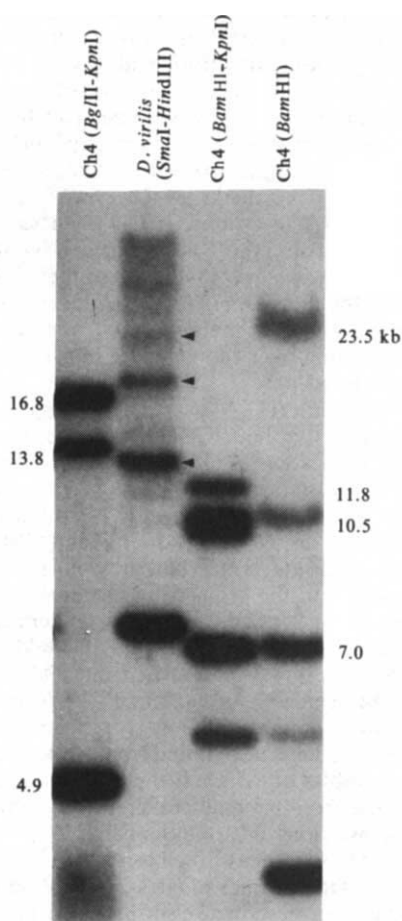
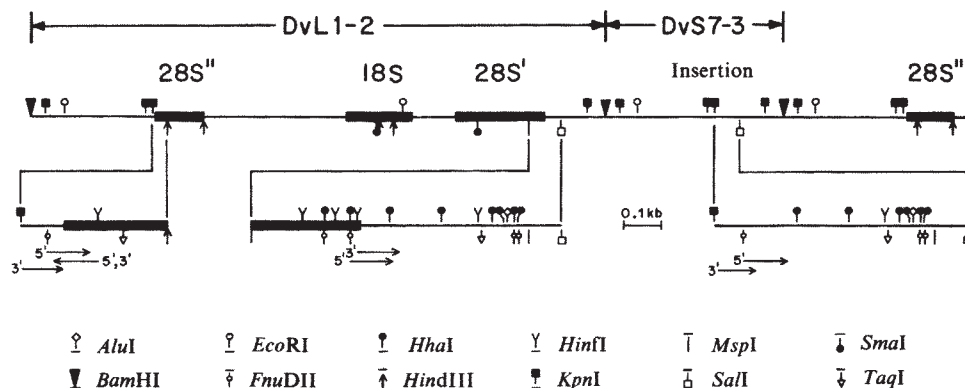


Fig. 2 Hybridization of the rDNA insertion segment DvS7-3 (Fig. 1) to a blot of *D. virilis* nuclear DNA digested with *Sma*I and *Hind*III and electrophoresed in a 0.4% agarose gel. Conditions have been described elsewhere⁸. With respect to insertions in the 28S coding region, the nearest *Sma*I site is ~2.2 kb upstream in the 28S' portion, and the nearest *Hind*III site is 278 bp into the 28S'' portion³. Lanes flanking the *D. virilis* lane contain markers of end-labelled segments of Charon 4 DNA^{5,16}: the first lane shows a *Kpn*I digest of Charon 4 (Ch4) end-labelled at *Bgl*II sites; the third lane shows a *Kpn*I digest of Charon 4 end-labelled at *Bam*HI sites; fourth lane, *Bam*HI segments of Charon 4. The sizes of some segments are indicated in kilobases¹⁶. Before blotting, DNA in the gel was fragmented by depurination to facilitate transfer of high-molecular weight DNA to the nitrocellulose filter¹⁷. Arrowheads indicate bands that are ~5-kb multiples of the darkest band at 7.5 kb.

by arrowheads in Fig. 2, are also 5-kb multiples of the 7.5-kb band, and the ladder may extend beyond the measurable range of segment sizes.

I determined the sequence of the very centre of a duplication to determine the precision of a tandem array. Figure 3 (centre) shows the sequence within the 0.7-kb *Kpn*I/*Sal*I segment of DvS7-3 that spans the end of one 5-kb insertion and the beginning of the next (Fig. 1). Also shown for comparison are the termini of insertions in the clone DvL1-2 (ref. 3). These are exactly represented in DvS7-3. Furthermore, they are separated in DvS7-3 by a 14-bp sequence, TGTCCTATCTACT, that flanks insertions as a direct repeat, and is present as a single copy in an uninterrupted unit of rDNA^{3,5}.

The data indicate that tandem repeats of a 5-kb rDNA insertion arose from unequal crossing-over in which the 14-bp flanking sequence served as the point of sequence homology. An event generating a 10-kb insertion from such recombination between two repeats having 5-kb insertions is shown in Fig. 4.

The reciprocal of the product with a 10-kb insertion is an rDNA repeat having no insertion (Fig. 4), and presumably recombination between intact genes and interrupted ones can also have occurred (this is the reverse of the event shown in Fig. 4). Thus it is possible that there is a flux between populations of rDNA repeats having 28S coding-region insertions and rDNA repeats that are intact. This has a bearing on an enigmatic feature of *Drosophila* rDNA: while repeats with insertions are inactive in rRNA synthesis⁹, and the insertions themselves are probably ancient and are greatly divergent among *Drosophila* species², coding-region sequences that flank insertions are largely indistinguishable from sequences in intact, functional genes^{3,5}. This is unexpected for gene copies that appear to have been nonfunctional for a very long time, particularly when insertion sequences immediately adjacent to coding regions are evidently under little or no pressure for sequence conservation^{2,5}. However, if coding regions in interrupted genes can be dissociated from insertions through unequal crossing-over involving the duplicated coding sequence at the flanks of insertions, and can exchange sequences with functional genes, then sequence integrity in the coding regions of interrupted genes would have been maintained.

There is one known case in which coding sequences adjacent to an insertion are not intact, that of the 5-kb type 1 insertion in *D. melanogaster* rDNA. There is a 23-bp deletion of coding sequence at the 28S' end of such insertions^{5,6}, and this deletion includes the 14 bp that are directly repeated at the ends of other insertions. Associated with the deletion are some other defects in the 28S' portion of the coding region⁵. One rDNA clone of three that have been examined in detail has another upstream deletion of 49 bp, and the results of blot hybridizations involving unfractionated rDNA have indicated the existence of other aberrations in the 28S' region of genes with

(28S' coding region)
 DvL1-2 TAGTGACGGCATGAATGATTACGAGATTCCTACTGTCCTATCTACTAGTTGCTTTCAGACAGTCGTTGGGAACAGACGTGT 3'
 DvS7-3 CGCGGAGGTGCTGACGTTAAGCCCACTGACTTTTCATGTCCTATCTACTAGTTGCTTTCAGACAGTCGTTGGGAACAGACGTGT 3'
 DvL1-2 CGCGGAGGTGCTGACGTTAAGCCCACTGACTTTTCATGTCCTATCTACTAGTACGAAACACAGCCAGGGAACGGGCTTGA 3'
 (28S' coding region)

Fig. 3 A comparison of sequences at the ends and at the centre of a 10-kb rDNA insertion in *D. virilis*. The middle line is from a *KpnI/SalI* segment of DvS7-3 (Fig. 1) that covers the junction of tandem repeats in a 10-kb insertion. The top line shows the sequence containing the junction between the 28S' rRNA coding region and one insertion in DvL1-2, and the sequence in the bottom line covers the 28S' end of the other insertion (see Fig. 1). Single underlining indicates the termini of insertions, and double underlining represents the 14 bases of 28S rDNA that are duplicated at the flanks of *D. virilis* rDNA insertions. Such a direct repeat of coding sequence at the flanks of insertions relates the insertions to transposable elements³. The identity of sequences in the DvS7-3 segment with those of DvL1-2 junctions strongly suggests that the 10-kb tandem duplication resulted from unequal crossing-over involving the flanks of insertions (see Fig. 4).

5-kb insertions. The 14 bp sequence is intact at the 28S' end of 5-kb insertions, and no coding region defects have been associated with the 28S' coding region⁵. There are no observed tandem repeats of the 5-kb insertion as there are in *D. virilis* rDNA. This also reflects the absence of the TGTCCCTATCTACT sequence at the 28S' end of 5-kb *D. melanogaster* insertions, as without it, unequal crossing-over events such as that shown in Fig. 4 cannot occur.

A problem with generalization of the unequal crossing-over model arises when 1- and 0.5-kb type 1 insertions in *D. melanogaster* rDNA are considered. These shorter insertions are flanked by identical or similar copies of the 14-bp sequence⁴, yet there are no observed duplications or higher multiplications of either length of insertion (only one-half of the 1-kb insertion is homologous with the 0.5-kb insertion). There is as yet no explanation for this apparent inconsistency between interrupted rDNAs in *D. melanogaster* and *D. virilis*.

Unequal crossing-over in *Drosophila* rDNA has been invoked to account for differences in the degree of the *bobbed* phenotype between certain parental strains and their progeny¹⁰; it has also been shown to occur meiotically between sister strands at the rDNA locus in yeast¹¹. The generation of tandem multiplications of a DNA segment containing the *ampC* (ampicillin resistance) gene in *Escherichia coli* has been suggested to involve recombination between copies of a 12-bp sequence that occurs twice within 10 kb of DNA that include *ampC*¹². Thus there are precedents for the model given in

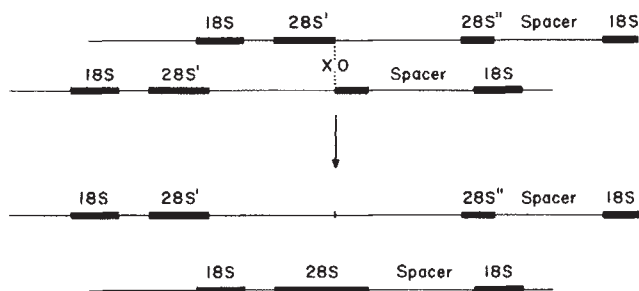


Fig. 4 Unequal crossing-over between *D. virilis* rDNA repeats with 5-kb insertions produces a repeat having a 10-kb insertion and one with an intact 28S rRNA coding region. The point of homology between repeats that is involved in recombination (dotted vertical line) is the 14-bp sequence directly repeated at the flanks of insertions. Crossing-over within this tetracaidecanucleotide generates a copy in the middle of the 10-kb tandem duplication, and this is indicated by a vertical dash. Sequence contexts are shown in Fig. 3.

Fig. 4. However, this is the first indication of rescue of coding sequences from an interrupted and transcriptionally inactive state to one in which the sequences can be functional in intact genes.

I thank Bruce Kohorn, Janice Oles, Anne Scott and Robert Wade for assistance. This work was supported by NIH grant GM 26378.

Received 1 December 1981; accepted 5 February 1982.

- Long, E. O. & Dawid, I. B. *Rev. Biochem.* **49**, 727-764 (1980).
- Barnett, T. & Rae, P. M. *Cell* **16**, 763-775 (1979).
- Rae, P. M. M., Kohorn, B. D. & Wade, R. P. *Nucleic Acids Res.* **8**, 3491-3504 (1980).
- Dawid, I. B. & Rebbert, M. L. *Nucleic Acids Res.* **9**, 5011-5020 (1981).
- Rae, P. M. M. *Nucleic Acids Res.* **9**, 4997-5010 (1981).
- Roiha, H., Miller, J. R., Woods, L. C. & Glover, D. M. *Nature* **290**, 749-753 (1981).
- Rae, P. M. M. & Steele, R. E. *Nucleic Acids Res.* **6**, 2987-2995 (1979).
- Rae, P. M. M., Barnett, T. & Murtif, V. M. *Chromosoma* **82**, 637-655 (1981).
- Long, E. O. & Dawid, I. B. *Cell* **18**, 1185-1196 (1979).
- Schalet, A. *Genetics* **63**, 133-153 (1969).
- Petes, T. D. *Cell* **19**, 765-774 (1980).
- Edlund, T. & Normark, S. *Nature* **292**, 269-271 (1981).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-530 (1980).
- Smith, M. et al. *Cell* **16**, 753-761 (1979).
- Tu, C.-P. D. & Cohen, S. N. *Gene* **10**, 177-183 (1980).
- DeWet, J. R. et al. *J. Virol.* **33**, 401-410 (1980).
- Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683-3687 (1979).

Flip-flop hydrogen bonding in a partially disordered system

W. Saenger* & Ch. Betzel

Abteilung Chemie, Max-Planck-Institut für experimentelle Medizin, D-3400 Göttingen, FRG

B. Hingerty & G. M. Brown

Divisions of Biology and Chemistry, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, USA

Cyclodextrins have proved useful as model systems for the study of hydrogen bonding¹. α -Cyclodextrin (α -CD) crystallizes from water as a hexahydrate having well defined hydrogen bonds (O—H...O) in linear and circular arrangements favouring 'endless' ...O—H...O—H...O—H... chains¹⁻³. The larger β -cyclodextrin (β -CD) forms a crystalline dodecahydrate, β -CD·12H₂O in which X-ray localization of the hydroxyl hydrogen atoms is made difficult because 25 of the 45 hydroxyl groups are statistically disordered and water molecules enclosed within the β -CD cavity are distributed over several sites that are not fully occupied⁴. We have therefore carried out a neutron diffraction study of β -CD at the Oak Ridge high flux isotope reactor to determine whether there is any well defined hydrogen bonding in disordered systems. In β -CD there are 19 hydrogen bonds of the type O—H...H—O. In these bonds, oxygen atoms are in the normal O—H...O distance range⁵, but two statistically half-occupied H atoms are arranged between them. The fact that the H...H separation of ~1 Å is so short that the two H atoms positions are mutually exclusive suggests an equilibrium between two states: O—H...O $\rightleftharpoons$ O...H—O. Of the two H atoms only one is in hydrogen-bonding contact at a given time; the other one is flipped out to form a hydrogen bond with an adjacent acceptor group and vice versa. Because long hydrogen-bonding chains are involved in a cooperative, concerted motion (domino effect), we have coined the term 'flip-flop hydrogen bond'. This study demonstrates that hydrogen bonds are operative in disordered systems and display dynamics even in solid state.

* To whom correspondence should be addressed. Present address: Institut für Kristallographie, Freie Universität Berlin, Takustrasse 6, D-1000 Berlin 33, FRG.

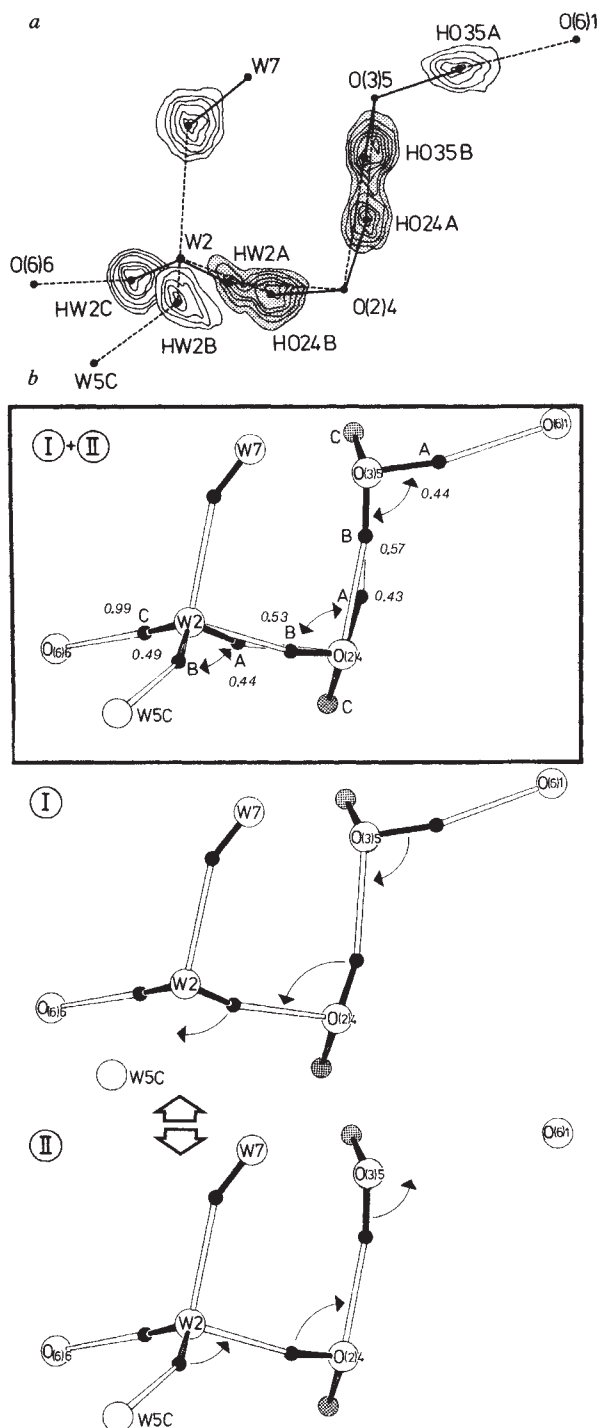


Fig. 1 *a*, Section of a Fourier difference map calculated from neutron diffraction data, showing D atom positions (referred to as H atoms in the text). Flip-flop hydrogen bonds $\text{O}-\text{H}\cdots\text{H}-\text{O}$ are indicated by stippling. $\bullet$, Positions of H, O and C atoms; —, covalent bonds; ---, hydrogen bonds. *b*, Same section as in *a*, giving an interpretation at the atomic level and showing the deconvolution of the flip-flop hydrogen bonding system into two states (I and II) having normal hydrogen bonds of the type $\text{O}-\text{H}\cdots\text{O}$. Solid arrows indicate the manner in which H atoms may jump in concerted, cooperative motion from one state to another. Circles of increasing size represent H, C and O atoms, with C shown shaded and H filled. Covalent bonds are indicated by solid lines, H bonds by open lines. The water molecule, W2, displays three H atom positions, one (C) fully occupied with an occupancy factor 0.99 (1) whereas A and B are only partly filled [0.44 (1) and 0.49 (1)]. In the text, H atoms A, B and C of W2 are designated HW2A, HW2B and HW2C and those attached to hydroxyl O(2)4 are termed HO24A and HO24B. All geometrical data are close to 'standard' values and will be published in detail elsewhere.

As $\beta\text{-CD}\cdot 12\text{H}_2\text{O}$ was crystallized from D_2O , replaceable hydroxyl and water H atoms were substituted by D. Here, however, we will only use the term H atom.

The hydrogen bonding scheme of $\beta\text{-CD}$ was assigned from 7,378 neutron diffraction data covering a resolution of 0.56 Å. Within one asymmetric unit of space group $\text{P}2_1$ containing $\beta\text{-CD}\cdot 12\text{H}_2\text{O}$, there are 39 hydrogen bonds of the type $\text{O}-\text{H}\cdots\text{O}$, and 19 of the type $\text{O}-\text{H}\cdots\text{H}-\text{O}$ in which two H atoms are located between oxygen atoms in a near-linear arrangement (Fig. 1*a*). In the latter, $\text{O}\cdots\text{O}$ separations are in the normal range (2.75–2.90 Å) as are covalent $\text{O}-\text{H}$ distances of ~ 1 Å. The apparent $\text{H}\cdots\text{H}$ contacts of between 0.86 and 1.18 Å, however, are much shorter than the 2.4 Å calculated from the sum of van der Waals' radii of the two H atoms⁶. The two H atom positions are therefore mutually exclusive and to prevent clashing, their positions are statistically only half filled and in each $\text{O}-\text{H}\cdots\text{H}-\text{O}$ hydrogen bond, H atom occupancies add up to ~ 1.0 . Therefore, if one of the two H atoms is present in the $\text{O}-\text{H}\cdots\text{H}-\text{O}$ hydrogen bond, for geometrical reasons the other one is pushed out to form a hydrogen bond with another partner as acceptor, and vice versa. This scheme was observed in all the 19 $\text{O}-\text{H}\cdots\text{H}-\text{O}$ bonds. Because, on a statistical basis, the oxygen atoms in such a 'flip-flop' system are always involved in hydrogen bonding, their positions and vibrational parameters are hardly affected by changes from one hydrogen bonding direction to another. Rather, they act as 'fixed centres' around which hydrogen atoms arrange or 'flip-flop' according to optimum bonding patterns.

A comparable scheme has been found previously in deuterated ice^{7,8} and in ice clathrates⁹. Two deuterium atoms are located between hydrogen-bonded oxygen atoms and in ice, they are responsible for the observed residual entropy of 0.87 entropy units (ref. 10). The salient difference between ice and the system considered here, however, is that in the high-symmetry (hexagonal and cubic) space groups of ice and ice clathrate, all O and D positions are constrained by symmetry requirements, whereas they are free to arrange in $\beta\text{-CD}\cdot 12\text{H}_2\text{O}$.

Figures 1 and 2 show examples of flip-flop hydrogen bonds. In Fig. 1, there are two flip-flop $\text{O}-\text{H}\cdots\text{H}-\text{O}$ hydrogen bonds connected by hydroxyl $\text{O}(2)4-\text{H}$, that is, $\text{W2}-\text{HW2A}\cdots\text{HO24B}-\text{O}(2)4$ and $\text{O}(2)4-\text{HO24A}\cdots\text{HO35B}-\text{O}(3)5$. The H atom of $\text{O}(2)4$ is twofold disordered, producing HO24A and HO24B with occupation factors of 0.43 (1) and 0.53 (1). Both positions cannot be filled simultaneously. Assuming HO24A to be occupied, then the hydrogen bonding chain (I) shown in Fig. 1*b* is formed involving all atoms indicated by A. If, on the other hand, position HO24A is empty and HO24B occupied, then a concerted change of H atoms and hydrogen bonds occurs so that all positions indicated by B are filled (II in Fig. 1*b*).

The water molecule W2 is engaged in two fully established hydrogen bonds, accepting from HW7A and donating to O(6)6, and therefore positions W2 and HW2C are fully occupied. The other hydrogen atom of W2 is part of the flip-flop system and disordered over HW2A and HW2B. All three hydrogen occupancies associated with W2 add up to 1.92 (2), in close agreement with the ideal 2.0. The two occupation factors of HW2B and HW2A, 0.44 (1) and 0.49 (1) add up to only 0.93 instead of the presumed 1.0. This is probably due to errors in measurement and refinement or is caused by incomplete H/D exchange.

In the more complicated scheme illustrated in Fig. 2, three flip-flop hydrogen bonds cluster around water molecule W1. The oxygen atom position of the latter is fully occupied, but its hydrogen atoms are statistically disordered over four tetrahedrally arranged sites with occupation factors in the range 0.34 (2)–0.60 (1), adding up to 1.93 (2). The arrangement of the water H atoms could correspond to statistical distribution over six different water orientations with unknown individual contributions.

Because the H atom positions in the flip-flop hydrogen bonds shown in Figs 1 and 2 are approximately half occupied with

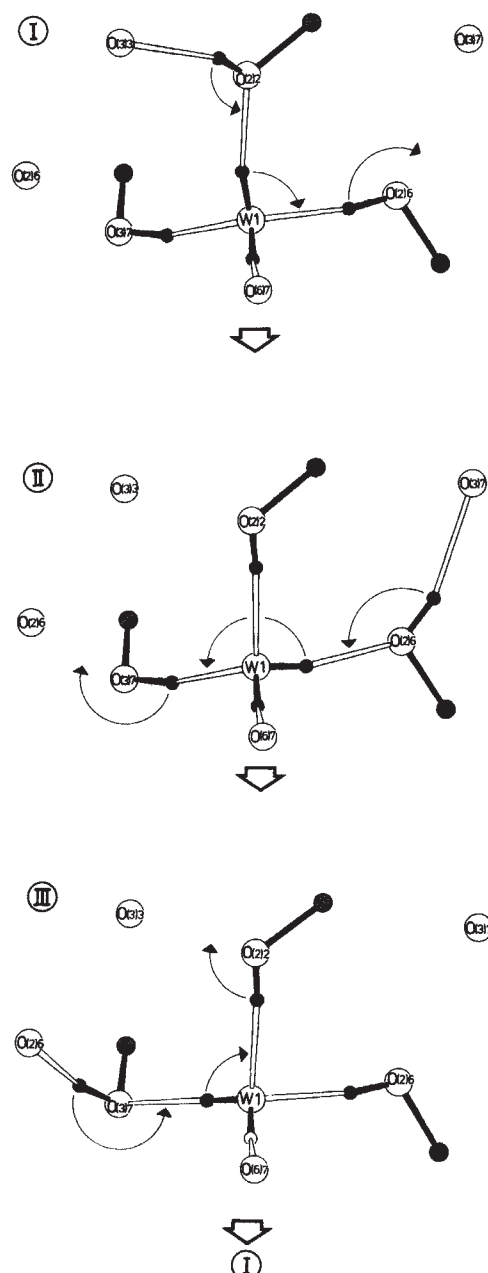
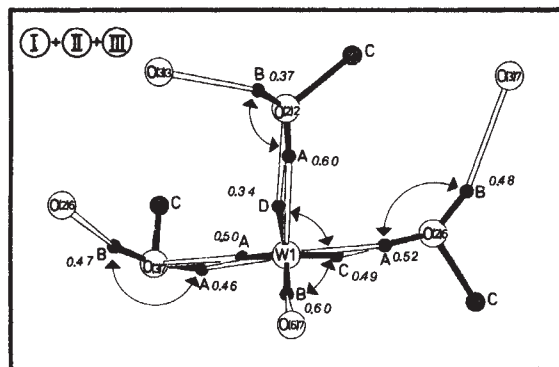
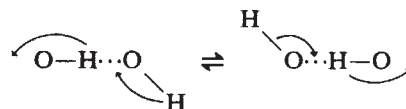


Fig. 2 A more complex system of three flip-flop hydrogen bonds connected by a water molecule, W1. The scheme observed (top) is separated into three schemes (I, II and III) having standard hydrogen bonds $\text{O}-\text{H}\cdots\text{O}$; more schemes are possible because W1 is probably oriented statistically into six different positions.

occupancy factors adding up to ~ 1.0 in each case, the two possible hydrogen bonds $\text{O}-\text{H}\cdots\text{O}$ and $\text{O}\cdots\text{H}-\text{O}$ are energetically comparable. This means that there may be transitions from one to the other, a jump of a H atom from one position to another causing all the connected hydrogen bonds to change orientation in a cooperative, concerted mechanism, reminiscent of a relay system (domino effect; see arrows in Figs 1, 2).

The crystallographically observed atomic arrangement $\text{O}-\text{H}\cdots\text{O}$ could be due either to averaging of crystal domains or of unit cells containing hydrogen bonds $\text{O}-\text{H}\cdots\text{O}$ and $\text{O}\cdots\text{H}-\text{O}$, or to direct observation of a dynamical equilibrium which also appears to prevail in ice^{9,10}.



Because we do not know the residual entropy of crystalline $\beta\text{-CD}\cdot 12\text{H}_2\text{O}$, we cannot assess which of the two possibilities is more likely. In solution, however, it is certain that the dynamical system will be present. Moreover, if single 'standard' hydrogen bonds $\text{O}-\text{H}\cdots\text{O}$ are compared energetically with flip-flop hydrogen bonds, the latter will be preferred because of the expected advantageous entropic term resulting from statistical, near-equivalent distribution of H atoms between the oxygen atoms. In this context, it is of interest that differential scanning calorimetric measurements of $\beta\text{-CD}\cdot 12\text{H}_2\text{O}$ crystals showed a strong endothermic effect at -46°C , suggesting that below this temperature, an ordering of the whole crystal structure occurs into one or other of the flip-flop orientations¹¹.

Thus we conclude that hydrogen bonding occurs even in badly disordered systems. Note that water molecules acting as joints within a system of flip-flop hydrogen bonds click into well defined orientations dictated by optimum hydrogen-bonding interactions and are by no means rotating to produce spherically averaged structures. The hydrogen bonds seem to operate in even less stringent conditions than previously assumed. This leads to flickering clusters of water occurring either in bulk or in hydration shells where disordered hydroxyl groups are dominant. If we combine flip-flop hydrogen bonds with circular structures formed by hydroxyl groups as observed in the α -cyclodextrin hydrates^{1,2,12}, an interesting dynamical system is obtained which consists of circles of four to six (or more) water hydroxyl groups with all $\text{O}-\text{H}\cdots\text{O}-\text{H}\cdots\text{O}-\text{H}$ hydrogen bonds running in the same direction (homodromic, see ref. 2) and oscillating clockwise and anti-clockwise in a flip-flop manner. Such systems are favoured energetically due to the cooperative effect³ as well as to the entropic contributions of two equivalent or near-equivalent states. In water clusters, such flip-flop circles can be interconnected to form larger, rapidly interconverting and changing systems.

G.M.B. was supported by the Division of Materials Sciences, Office of Basic Energy Sciences, US Department of Energy, under contract W-7405-ENG-26 with the Union Carbide Corporation. B.H. was supported by the Office of Health and Environmental Research, US Department of Energy, under contract W-7405-Eng-26 with the Union Carbide Corporation.

Received 27 August 1981; accepted 22 February 1982.

- Klar, B., Hingerty, B. & Saenger, W. *Acta crystallogr.* B36, 1154-1165 (1980).
- Saenger, W. *Nature* 279, 343-344 (1979).
- Lesyng, B. & Saenger, W. *Biochim. biophys. Acta* (in the press).
- Lindner, K. & Saenger, W. *Carbohydrate Res.* (in the press).
- Jeffrey, G. A. & Tagaki, S. *Acc. chem. Res.* 11, 264-270 (1978).
- Weast, R. C. (ed.) *CRC Handbook of Chemistry and Physics*, D-178 (CRC Press, Cleveland, 1976).
- Peterson, S. W. & Levy, H. A. *Acta crystallogr.* 10, 70-76 (1957).
- Kuhs, W. F. & Lehmann, M. L. *Nature* 294, 432-436 (1981).
- Hollander, F. & Jeffrey, G. A. *J. chem. Phys.* 66, 4699-4705 (1977).
- Pauling, L. *J. Am. Chem. Soc.* 57, 2680-2684 (1935); *The Nature of Chemical Bond* 3rd edn, 467 (Cornell University Press, 1960).
- Yamazaki, M. *et al. Chem. Lett.* (in the press).
- Chacko, K. K. & Saenger, W. *J. Am. chem. Soc.* 103, 1708-1715 (1981).

Conversion of the chemical energy of methylmalonyl-CoA decarboxylation into a Na^+ gradient

Wilhelm Hilpert & Peter Dimroth

Institut für Physiologische Chemie der Technischen Universität München, Biedersteiner Strasse 29, D-8000 München 40, FRG

Primary active transport systems establish electrochemical gradients of their transport substrates at the direct expense of chemical or light energy. The chemical energy can be provided by cleavage of the phosphoric anhydride bond of ATP, establishing Na^+ , K^+ , H^+ or Ca^{2+} gradients with the respective ATPases. Whereas in eukaryotic cells the Na^+ and K^+ imbalance between cells and the surrounding fluids is brought about by the $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ ¹, no such enzyme is known for bacterial cells. Generally, the gradients of Na^+ and K^+ ions in these cells are believed to be generated by secondary transport systems consuming the energy provided by the electrochemical gradient of protons^{2,3}. In some bacterial species, however, sodium ions may be transported by primary ATP or light-driven sodium pumps^{4,5}. The primary sodium pump oxaloacetate decarboxylase represents a novel energy conversion mechanism in that it consumes the chemical energy of the decarboxylation of oxaloacetate to drive Na^+ transport⁶⁻⁸. We show here that methylmalonyl-CoA decarboxylase is another example of a Na^+ transport enzyme converting the chemical energy of a decarboxylation reaction into an ion gradient.

The action of oxaloacetate decarboxylase in catalysing the transport of Na^+ ions across the cell membrane, thereby generating an electrochemical gradient of Na^+ ions⁶⁻⁸, is in accord with its localization in the cell membrane and the specific requirement of Na^+ ions for catalytic activity⁹⁻¹¹. By analogy, we hypothesized that methylmalonyl-CoA decarboxylase might also function in this way. As this enzyme apparently did not require a specific metal ion for catalytic activity, it could catalyse the transport of protons.

Methylmalonyl-CoA decarboxylase from *Micrococcus laticyticus* (syn. *Veillonella alcalescens*) is reported to be tightly bound to the ribosomes^{12,13}, a property not compatible with a transport function for this enzyme. We have therefore reexamined the localization of the decarboxylase using chromatography of the particulate cell fraction on Sephacryl S-1000 Superfine columns. Ribosomes and membranes were clearly resolved, as indicated from the separation of the markers succinate dehydrogenase (for the membrane) and ribose (for ribosomes) and also from the change in the ratio of absorbances at 280 and 260 nm (~ 1 in the membrane and ~ 0.5 in the ribosomal fractions). The distribution of methylmalonyl-CoA decarboxylase indicated that it was attached to the membranes but not to ribosomes.

The enzyme was solubilized from the membranes with Triton X-100 and purified about 30-fold over the crude membrane extract on an avidin-Sepharose affinity column¹⁴. Examination of the cation requirement of the purified enzyme indicated that the decarboxylase was specifically activated by Na^+ ions ($K_{\text{Mapp}} \sim 1 \text{ mM}$ at 0.10 mM (*R, S*)methylmalonyl-CoA). This property of the enzyme and its binding to the cell membrane are analogous to properties of oxaloacetate decarboxylase and are compatible with a function of the decarboxylase as a sodium pump.

Sodium transport was determined from the amount of radioactivity taken up into inverted vesicles of *V. alcalescens* on incubation with $^{22}\text{NaCl}$. Whereas in the absence of methyl-

malonyl-CoA almost no Na^+ was taken up by the vesicles, substrate addition resulted in the rapid accumulation of Na^+ ($\sim 1.1 \text{ nmol } ^{22}\text{Na}^+$ per mg protein within the first 15 s; Fig. 1). Due to the high activity of the decarboxylase ($0.35 \text{ U per mg protein}$), all methylmalonyl-CoA was decarboxylated within the 15 s, thus terminating Na^+ uptake. The accumulated Na^+ was not retained in the vesicles but returned to the medium at an initial rate of $\sim 0.55 \text{ nmol per min per mg protein}$, probably due to leakiness of the membranes. These observations are completely analogous to the kinetics of Na^+ transport into vesicles from *Klebsiella aerogenes* as catalysed by oxaloacetate decarboxylase⁷.

The action of methylmalonyl-CoA decarboxylase as a Na^+ transport enzyme was confirmed by treating the vesicles with avidin, which simultaneously abolished Na^+ transport and

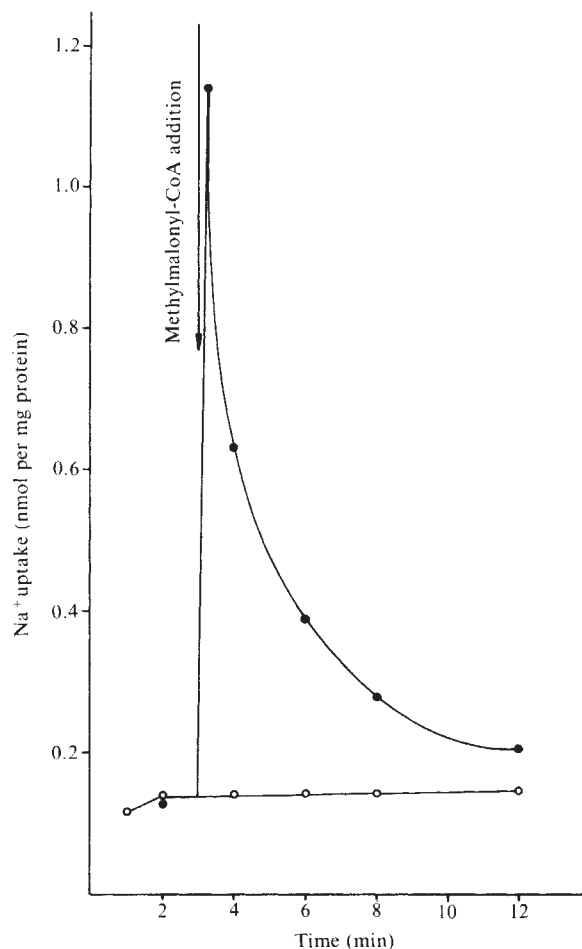


Fig. 1 Kinetics of $^{22}\text{Na}^+$ uptake into inverted membrane vesicles from *V. alcalescens*. The vesicles were prepared by passing *V. alcalescens* cells in $50 \text{ mM K-phosphate buffer, pH 7.5}$, through a French press at $8,000 \text{ p.s.i.}$ and were subsequently isolated by centrifugation steps as described previously⁷. The inverted vesicles ($3.4 \text{ mg protein, } 0.35 \text{ U per mg methylmalonyl-CoA decarboxylase}$) were incubated at 25°C in a total volume of $10.67 \text{ ml } 50 \text{ mM K-phosphate buffer, pH 7.5}$, with $0.44 \text{ mM } ^{22}\text{NaCl}$ ($2,800 \text{ counts per min per nmol}$). At the arrow, $0.3 \text{ mM (R, S)methylmalonyl-CoA (Li-Salt)}$ was added and $^{22}\text{Na}^+$ uptake was determined from samples (0.1 ml) taken at appropriate incubation periods and passed over columns ($0.5 \times 2 \text{ cm}$) of Dowex 50, K^+ to separate free Na^+ from Na^+ trapped inside the vesicles. The radioactivity of internal $^{22}\text{Na}^+$ was subsequently determined by liquid scintillation counting (●). ○, Control without methylmalonyl-CoA addition.

Table 1 Avidin-induced inhibition of methylmalonyl-CoA decarboxylase-catalysed Na⁺ uptake into inverted vesicles

Pretreatment of vesicles	Methylmalonyl-CoA decarboxylase activity (U per mg protein)	²² Na ⁺ uptake after incubation periods of	
		0.25 min	1 min
None	0.34	1.14	0.63
Avidin-biotin complex	0.36	1.09	0.65
Avidin	0	0.11	0.15
None, methylmalonyl-CoA omitted	0	0.13	0.14

See Fig. 1 legend for experimental conditions.

decarboxylase activity (Table 1). A parallel incubation with an avidin-biotin complex, however, did not affect either of these activities, ruling out the possibility that the avidin might act on the transport system in any way other than by a complexing of the biotin prosthetic group of the decarboxylase. The products of the decarboxylation, propionyl-CoA and bicarbonate, were unable to substitute for methylmalonyl-CoA in inducing Na⁺ accumulation to any considerable extent, thus confirming that Na⁺ is transported during the decarboxylation reaction. The function of methylmalonyl-CoA decarboxylase as a Na⁺ transport enzyme was further confirmed by reconstitution experiments. When the decarboxylase was purified by affinity chromatography on avidin-Sepharose and incorporated into phospholipid vesicles by the detergent dilution method¹⁵, using octylglucoside as the detergent, the reconstituted proteoliposomes were able to concentrate Na⁺ ions about six times over the incubation medium in response to the decarboxylation of methylmalonyl-CoA.

The mechanism of the Na⁺ transport was investigated by analysing the effect of certain ionophores. If the Na⁺ transport was electrogenic, creating a membrane potential, this could be limiting for the amount of Na⁺ accumulation, as has been shown to be the case for oxaloacetate decarboxylase⁷. The membrane potential would be discharged in the presence of valinomycin and K⁺ by the electrophoretic movement of K⁺ ions. In these conditions, about twice as much Na⁺ was accumulated in the vesicles as in the valinomycin-free controls (Table 2), thereby indicating the presence of electrogenic Na⁺ transport. The uncoupler carbonylcyanide-*p*-trifluoromethoxy phenylhydrazone also increased Na⁺ uptake within the vesicles, although less efficiently. This effect is probably also due to discharging of the membrane potential. Furthermore, the result indicates that the generation of a proton gradient is not the primary event for the Na⁺ transport. As expected, the accumulation of Na⁺ inside the vesicles was completely abolished by the presence of the Na⁺ ionophore nigericin.

Thus, our results characterize methylmalonyl-CoA decarboxylase as a new transport enzyme for Na⁺ ions. The enzyme

Table 2 Effect of ionophores on ²²Na⁺ uptake into inverted membrane vesicles

Addition	²² Na ⁺ uptake after incubation for	
	0.25 min	1 min
	(nmol per mg protein)	
None	1.14	0.63
Valinomycin, 22 µM	2.22	1.24
CCFP, 20 µM	1.56	0.65
Nigericin, 22 µM	0.14	0.14
None, methylmalonyl-CoA omitted	0.13	0.14

See Fig. 1 legend for experimental conditions. CCFP, carbonylcyanide-*p*-trifluoromethoxy phenylhydrazone.

is closely related to the Na⁺ pump oxaloacetate decarboxylase⁶⁻⁸ and both enzymes constitute a new group of vectorial catalysts, converting the energy liberated by decarboxylation reactions into ion gradients. It is well known that the decarboxylation of certain 'energy-rich' carbonic acids is used as a driving force in biosynthesis, for example, in the biosynthesis of fatty acids and of phosphoenolpyruvate, where the decarboxylation of malonyl-CoA and oxaloacetate, respectively, shifts the equilibrium towards synthesis. The conversion of the energy of decarboxylation reactions into Na⁺ gradients by methylmalonyl-CoA decarboxylase and oxaloacetate decarboxylase, however, is a previously unknown biological use of decarboxylation energy, one closely related to other energy converting processes such as the formation of cation gradients by ATP hydrolysis.

This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

Received 19 January; accepted 25 February 1982.

- Glynn, I. M. & Karlsh, S. J. D. A. *Rev. Physiol.* **37**, 13-55 (1975).
- Mitchell, P. J. *Bioenerg.* **4**, 53-91 (1973).
- Rosen, B. P. & Kashket, E. R. in *Bacterial Transport* (ed. Rosen, B.) 559-620 (Dekker, New York, 1978).
- Heefner, D. L., Kobayashi, H. & Harold, F. M. *J. biol. Chem.* **255**, 11403-11407 (1980).
- McDonald, R. E., Greene, R. V., Clark, R. D. & Lindley, E. V. *J. biol. Chem.* **254**, 11831-11838 (1979).
- Dimroth, P. *FEBS Lett.* **122**, 234-236 (1980).
- Dimroth, P. *Eur. J. Biochem.* **121**, 443-449 (1982).
- Dimroth, P. *J. biol. Chem.* **256**, 11974-11976 (1981).
- Stern, J. R. *Biochemistry* **6**, 3545-3551 (1967).
- Dimroth, P. *Eur. J. Biochem.* **115**, 353-358 (1981).
- Dimroth, P. *Eur. J. Biochem.* **121**, 435-441 (1982).
- Galivan, J. H. & Allen, S. H. G. *J. biol. Chem.* **243**, 1253-1261 (1968).
- Galivan, J. H. & Allen, S. H. G. *Archs Biochem. Biophys.* **126**, 838-847 (1968).
- Henrikson, K. P., Allen, S. H. G. & Maloy, W. L. *Analyt. Biochem.* **94**, 366-370 (1979).
- Racker, E., Violand, B., O'Neal, S., Alfonzo, M. & Telford, J. *Archs Biochem. Biophys.* **198**, 470-477 (1979).

Erratum

The article by F. W. Alt, N. Rosenberg, R. J. Casanova, E. Thomas and D. Baltimore, *Nature* **296**, 325-331, was given the incorrect title. It should read 'Immunoglobulin heavy chain class switching and inducible expression in an Abelson murine leukaemia virus transformed cell line'.

Corrigenda

In the article 'Molecular cloning establishes proenkephalin as precursor of enkephalin-containing peptide' by U. Gubler *et al.*, *Nature* **295**, 206-208 (1982), in Fig. 2 the residue at position 20, shown as Asp, should be Asn, and the sequence around it Thr-Asp-Leu-Asn-Pro-Leu.

In the article 'The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences' by G. E. Bell *et al.*, *Nature* **295**, 31-35 (1982), the first three lines in Table 2 should read:

Sequence	Occurrences	
	Allele 1 (λHi1)	Allele 2 (λHi2)
a, ACAGGGGTGTGGGG	15 (44.1%)	24 (53.3%)
b, ACAGGGGTCTGGGG	5 (14.7%)	6 (13.3%)
c, ACAGGGGTCTGGGG	7 (20.5%)	7 (15.6%)

and the last two lines:

Arrangement of oligonucleotides

Allele 1 5' cdi aba aba dab aab aca aaa cea faa cgh ccc b 3' (34 repeats)

Allele 2 5' cdi ifa faa aba baa aaf aba afa baa aac aaa cac baa afc ccb 3' (45 repeats)

MATTERS ARISING

Thermomechanical inferences from deep seismic sounding sections

CHOWDHURY AND HARGRAVES¹ recently interpreted a deep seismic sounding section, between Kavali and Udipi in India, which appears to show deep faults offsetting an essentially horizontal Mohorovičić discontinuity. Reflectors below the Moho seem to have lower and more uniform apparent dips than those above it, leading these authors to suggest that "at the time of its formation [before faulting] the Moho constituted a thermomechanical boundary between rigid [folding] crust and plastic [ductilely flowing] mantle"¹. Although elements of this interpretation are questionable (notably that folding constitutes rigid behaviour), their subsequent conclusions ignore even more fundamental considerations regarding the rheology of solids: "Ductile flow requires that the then Moho level temperatures were close to the crustal solidus [of 1,000 °C for anhydrous granulite³]. Consequently, mean geothermal gradients must have been in the range 20–30 °C km⁻¹, two or three times those inferred for present day shields^{4,11}."

The mere fact that a rock has been plastically deformed does not uniquely determine the temperature at which that deformation took place. Steady homogeneous flow in a solid is governed by a constitutive equation relating temperature, strain rate, and stress. For a rock subjected to uniaxial compression the equation relating the shortening rate $\dot{\epsilon}$ to the differential stress σ and absolute temperature T is typically of the form

$$\dot{\epsilon} = A \exp(-Q/RT)f(\sigma) \quad (1)$$

where Q is an activation energy, R is the universal gas constant and A is a constant. The form of the function f depends on the dominant deformation mechanism(s), whereas the constants A and Q are determined by the minerals present, their grain size and geometry, and the activities of any volatile phases.

Solving equation (1) for the temperature gives

$$T = Q\{R \ln[Af(\sigma)/\dot{\epsilon}]\}^{-1} \quad (2)$$

from which it is evident that the temperature required for flow can be reduced either by decreasing the strain rate or by increasing the stress. The validity of this conclusion is not limited to the case of uniaxial deformation; equations (1) and (2) can be extended to more general states of deformation, although the measures of strain rate and stress must be

appropriately generalized and the function $f(\sigma)$ may take a different form.

To determine unambiguously the temperature at which plastic deformation ceased (and at which the Moho formed in the Chowdhury-Hargraves model) we would need to know: (1) whether the deformation was homogeneous and steady; (2) what was the appropriate constitutive relation; and (3) what were the values of $\dot{\epsilon}$ and σ just below the Moho when it formed. Condition (1) may be

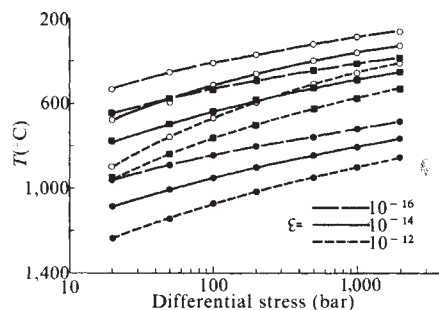


Fig. 1 Flow temperature as a function of differential stress for three strain rates. ○, Quartz⁶; □, diabase⁷; ●, olivine⁵.

approximately satisfied for deep processes and geological time scales but requirements (2) and (3) cannot easily be met. If simplifying assumptions could be made about sub-Moho mineralogy and deformation mechanisms, an empirical flow law⁵⁻⁹ might be used. Similarly it could be assumed that currently estimated deep crustal or mantle flow stresses (several tens of bars to several kilobars)¹⁰ and strain rates (10^{-17} – 10^{-12} s⁻¹)^{11,12} were applicable to the process of Moho formation. Using flow laws for quartz (ref. 6 and unpublished data), diopside⁸ and olivine⁹ aggregates along with the extreme values of stress and strain rate noted above, upper and lower bounds may be placed on flow temperature from equation (2). Table 1 shows these values together with their respective mean geothermal gradients for crustal thicknesses of 30–70 km.

Evidently Chowdhury and Hargraves' model places no real constraints on crustal thermal structure; in fact their near crustal

solidus Moho formation temperature and corresponding geothermal gradients are nearly maximal values in a broad spectrum of possibilities. Unique determination of thermal structure by this method requires petrological and rheological constraints which are not generally available and cannot be assumed without prejudicing the results.

I thank W. M. Bruner and P. Koch for helpful suggestions.

P. S. KOCH

Department of Earth and Space Sciences,
University of California,
Los Angeles,
California 90024, USA

1. Chowdhury, K. R. & Hargraves, R. B. *Nature* **291**, 648–650 (1981).
2. Kaila, K. L. et al. *J. geol. Soc. Ind.* **20**, 307–333 (1979).
3. O'Hara, M. J. *J. geol. Soc. Lond.* **134**, 185–200 (1977).
4. Rao, R. U., Rao, G. V. & Narain, H. *Earth planet. Sci. Lett.* **30**, 57–64 (1976).
5. Christie, J. M., Koch, P. S. & George, R. P. Jr *EOS* **60**, 948–949 (1979).
6. Koch, P. S., Christie, J. M. & George, R. P. *EOS* **61**, 376 (1980).
7. Jaoul, O., Shelton, G. & Tullis, J. *EOS* **62**, 396 (1981).
8. Shelton, G. & Tullis, J. *EOS* **62**, 396 (1981).
9. Post, R. L. *Tectonophysics* **42**, 75–110 (1977).
10. Hanks, T. C. & Raleigh, C. B. *J. geophys. Res.* **85**, 6083–6085 (1980).
11. Post, R. L. thesis, Univ. California, Los Angeles (1973).
12. Heard, H. C. *Phil. Trans. R. Soc. A* **283**, 173–186 (1976).

ROY CHOWDHURY AND HARGRAVES REPLY—Koch's argument pertains to the validity of an indirect consequence of our conclusion—the inferred geothermal gradient in Archaean times—without disputing the seismic evidence which prompted our overall speculation: that in Archaean times, the base of the crust was the base of the lithosphere¹. Nevertheless his point is well taken, because we had not considered rheology in our paper.

In essence, Koch is stating that if one takes the extreme estimates of the relevant rheological variables, one gets a wider range in the resulting geothermal gradient for our model than the 20–30 °C km⁻¹ we had inferred. These variables are (1) crustal composition and thickness, (2) stress and strain rate, (3) flow law. We do not dispute this, but would rather ask: what physical conditions are required to permit our interpre-

tation of the seismic data in terms of the current understanding of the rheological properties of the crust and upper mantle?

It seems reasonable that the Archaean continental crust was differentiated much as it is today²—in which case anhydrous mafic granulites seem to be the likely material composing the lower crust. The “smearing out”, we mentioned, would require this material to be ductile at its base. We used 38 km for the crustal thickness because it is the present average for the region under study. A 70-km thick Archaean crust seems unlikely as the Dharwar schist belts now exposed on the surface are only in the greenschist facies.

We do not believe that the possible range of values for the differential stress considered by Koch (tens to thousands of bars) is the probable range. Recent studies³ tend to suggest low stresses (10–100 bar) associated with contemporary asthenospheric flow. Similarly 10^{-14} s^{-1} is considered to be a geologically representative strain rate⁴.

Unfortunately rheological behaviour of geological materials in the above mentioned conditions is also not known uniquely. For example, the flow law of Goetze⁵ for olivine, reported to be valid below 2 kbar, yields a flow temperature of 1,235 °C for $\sigma = 20$ bar and $\dot{\epsilon} = 10^{-12} \text{ s}^{-1}$ as compared with 1,040 °C from the law used by Koch. There is also the problem of extrapolating experimental values to low stresses and strain-rates.

Figure 1 shows the dependence of the flow temperature on differential stress for three compositions and three strain rates using flow laws for quartz⁶, diabase⁷ and olivine⁵. Clearly, the temperature is sensitive to chemical composition and the sensitivity increases as the differential stress decreases. Thus, if the smearing out of the base of the crust took place under relatively low stress fields (tens of bars), then the chemical composition becomes a particularly important factor in determining the actual temperature of the process. For a strain rate of 10^{-14} s^{-1} we obtain a temperature range between 780 °C for diabase and 1,080 °C for olivine at 20 bar. If, as we had inferred, the structural discordance at 38 km is taken to indicate that the mantle (olivine) below was ductile and convecting, then, in the abovementioned stress and strain rate conditions, a temperature of $\geq 1,080$ °C and a gradient ≥ 28 °C km⁻¹ are implied. As olivine appears stronger than other rock forming materials, this model requires that the lower crust must then have been ductile as well, and hence the base of the lithosphere must have been above the Moho—crust being defined in chemical terms without consideration of its rheological properties. If, on the other hand, the boundary marks that point in time when the descending brittle–ductile boundary in the crust reached its base, then in the same conditions, a temperature of 780 °C, and a gradient of

20 °C km⁻¹ is implied for a ‘diabasic’ type lower crust.

Note that the thermal gradients inferred above are entirely consistent with those claimed in our original paper, which were based on experimental evidence on the beginning of melting in silicate rocks⁸. With declining heat flow, model 1 would evolve to model 2, and in the interim, the lithosphere would include a ductile lower crust above a brittle upper mantle (“jelly sandwich”) (J. Suppe, personal communication). Although model 1 may well have applied in the early Archaean, model 2 seems to be more reasonably associated with the onset of cratonic stability—which marks the end of the Archaean.

Clearly evolution of tectonic style in the Earth is associated with declining heat flow. Consideration of lithospheric evolution in rheological terms, with progressive descent of the brittle–ductile boundaries in both crust and mantle, as the Earth cooled, may help to explain the more abrupt changes.

We thank Brian Evans for discussions and computational help.

K. ROY CHOWDHURY
R. B. HARGRAVES

Department of Geological and
Geophysical Sciences,
Princeton University,
New Jersey 08544, USA

1. Roy Chowdhury, K. & Hargraves, R. B. *Nature* **291**, 648–650 (1981).
2. Zartman, R. E. & Wasserburg, G. J. *Geochem. cosmochim. Acta* **33**, 901–942 (1969).
3. Stocker, R. L. & Ashby, M. F. *Rev. Geophys.* **11**, 391–497 (1973).
4. Mercier, J.-C. C., Anderson, D. A. & Carter, N. L. *Pageophysics* **115**, 199–226 (1977).
5. Goetze, C. *Phil. Trans. R. Soc. A* **288**, 99–119 (1978).
6. Brace, W. F. & Kohlstedt, D. L. *J. geophys. Res.* **85**, 6248–6252 (1980).
7. Caristan, Y. D. thesis, Massachusetts Institute of Technology (1980).
8. O'Hara, M. J. *J. geol. Soc. Lond.* **134**, 185–200 (1977).

Behaviour, paternity and testes size

In a recent issue of *Nature*^{1,2}, Martin and May, and Harcourt *et al.* proposed a theory that behaviour, paternity and testes size are related. This hypothesis is interesting but suffers from the use of, as a model, species in which sperm production per unit of testes, sperm per ejaculate and other important considerations can only be surmised. Perhaps larger testes size is a result of more frequent copulations or larger testes stimulate males to copulate more frequently. Are sperm numbers per ejaculate well correlated to testis size in primates? The boar, ram and bull have roughly the same testis size but daily sperm production per gramme of testis varies greatly between them, and the number of sperm per ejaculate is 10–50 times greater in the boar than in the bull or ram.

The time of mating relative to ovulation seems to be more important than any other one factor in determining paternity when two males are used at an interval in the rabbit, pig and sheep^{3,4}. When matings or inseminations of two males are coincident, then the fertility of the male and numbers of sperm determine paternity. With all these complex interactions playing an important part, it seems overly simplistic to ascribe much to testes size.

PHILIP DZIUK

Department of Animal Science,
College of Agriculture,
University of Illinois at
Urbana–Champaign,
Urbana, Illinois 61801, USA

1. Martin, R. D. & May, R. M. *Nature* **293**, 7–9 (1981).
2. Harcourt, A. H., Harvey, P. H., Larson, S. G. & Short, R. V. *Nature* **293**, 55–57 (1981).
3. Miller *et al.* *J. Reprod. Fert.* **19**, 545 (1969).
4. Dziuk, P. *J. Reprod. Fert.* **22**, 277 (1970).

HARCOURT REPLIES—I agree, as stated in our paper¹, that not only are more data needed to prove fully the idea that relative testes size is related to breeding system via sperm competition but also that other factors influence the correlation, and also agree that the association between relative testes size and sperm output is largely an assumption. Nevertheless, I would argue that the correspondence between the data and predictions based on this assumption is so good that unless another hypothesis explains equally well the association between relative testes size and breeding system, ours (with its assumption) must stand as the best explanation of the correlation.

Four points can be usefully commented on in more detail. (1) Although data are lacking on, for example, sperm numbers per ejaculate, they are not totally absent, and what information there is supports the hypothesis^{1,2}.

(2) Dziuk writes about testes size *per se*; we consider relative testes size, that is, testes weight per unit body weight. The difference is extremely important, and the fact that the boar, with a greater relative testes size than the bull (and I believe the ram?), produces more sperm per ejaculate than does the bull or ram fits our theory and its assumption.

(3) Timing of mating in relation to ovulation is of prime importance in determining paternity. However, primate males cannot judge precisely the time of ovulation. In this situation the male that inseminates the largest amounts of sperm over the longest periods will be at a competitive advantage when more than one male mates with the periovulatory female. Our argument is that to do this he requires a large volume of spermatogenic tissue and hence large testes. Dziuk implies just this

in his statement that "When matings... of two males are coincident, then... numbers of sperm determine paternity" (see, for example, ref. 3). Moreover, this statement made in relation to rabbit, pig and sheep supports our suggestion that the association between breeding system and relative testes size will be found in other mammalian orders.

(4) Finally, while the hypothesis relating breeding system to relative testes size (rather than to some more precise variable) is simple, the good agreement between the data and predictions based on the stated assumption implies that it is not simplistic.

A. H. HARCOURT

Department of Applied Biology,
University of Cambridge,
Pembroke Street,
Cambridge CB2 3DX, UK

1. Harcourt, A. H., Harvey, P. H., Larson, S. G. & Short, R. V. *Nature* **293**, 55–57 (1981).
2. Short, R. V. *Adv. Study Behav.* **9**, 131–158 (1979).
3. Dewsbury & Hartung *Anim. Behav.* **28**, 95–102 (1980).

Hexose transport in hybrids between malignant and normal cells

IT HAS been reported by White *et al.*¹ that a decreased K_m for hexose transport correlates with malignancy in matched pairs of hybrids between malignant and normal cells. Hexose transport was measured by the uptake of 0.1–5 mM 2-deoxy[³H]glucose at 20 °C; in these conditions, uptake was found to be linear for at least 10 min. No doubt the authors are aware that sugars such as 2-deoxyglucose (or D-glucose) enter cells rapidly, soon (generally within <30 s at 25 °C) establish a steady state—at intracellular concentrations which can approach that in the medium—and are, in turn, phosphorylated at slower rates which persist linearly for several minutes^{2,3}. Within 5 min of exposing Novikoff² or Lettrec⁴ cells (at 25 or 37 °C) to 2-deoxy[³H]glucose at concentrations of ≤0.2 mM, for example, >85% of the radioactivity associated with the cells is in the form of 2-deoxy[³H]glucose-6-phosphate, whereas <15% is present as an intracellular pool of free sugar.

In other words, it is likely that the rate-limiting step is phosphorylation, not transport, of the hexose and that what White *et al.* are measuring is the linear accumulation of 2-deoxy[³H]glucose-6-phosphate in cells. While the use of a 'coupled reaction' to measure a preceding one is a commonly used biochemical technique, its validation does depend on confirmation that the second reaction is not rate-limiting. The article by White *et*

Table 1 Kinetic parameters of 3-O-methyl-D-glucose uptake in malignant and non-malignant cells

Cell type	V_{max}	K_m
CBAT ₆ T ₆	325.1 ± 26.8	4.419 ± 0.520
fibroblasts	420.3 ± 90.8	4.868 ± 1.214
PG19	106.6 ± 5.1	2.114 ± 0.172
	116.6 ± 18.2	2.211 ± 0.460
1Acn2	331.2 ± 40.2	8.495 ± 1.282
1Acn1TG	107.6 ± 7.3	2.771 ± 0.241

The uptake of 3-O-methyl-D-glucose was measured by the technique previously described¹ for 2-deoxyglucose with the following modifications: (1) 3-O-methyl-D-[1-³H]glucose (0.1–5.0 mM; 10–500 mCi mmol⁻¹) was added to the cells in a volume of 250 µl and its uptake terminated by washing the cells twice with 1 ml of ice-cold phosphate-buffered saline. (2) At 20 °C, uptake of 3-O-methylglucose is linear for only 1–2 min, so that correspondingly shorter incubation times were necessary. Kinetic constants were calculated as before from the linear portions of the uptake plots.

*al.*¹, however, contains no mention of the phosphorylation reaction, let alone measurement of intracellular 2-deoxy[³H]glucose and 2-deoxy[³H]glucose-6-phosphate. Moreover, some authors have suggested that it is sugar kinases (not sugar transport *per se*) that are different in transformed and normal cells⁵.

Although the results of White *et al.*¹ clearly show that malignant cells take up hexoses at low concentrations better than their non-malignant homologues, the claim of these authors that their "findings leave little doubt that malignancy is closely linked to a decrease in the K_m of the membrane hexose transport system" seems as yet unjustified.

C. A. PASTERNAK
K. J. MICKLEM

Department of Biochemistry,
St George's Hospital Medical School,
London SW17 0RE, UK

K. R. F. ELLIOTT

Department of Biochemistry,
University of Manchester,
Manchester M13 9PL, UK

P. G. W. PLAGEMANN
R. M. WOHLHUETER

Department of Microbiology,
University of Minnesota,
Minneapolis, Minnesota 55455, USA

1. White, M. K., Bramwell, M. E. & Harris, H. *Nature* **294**, 232–235 (1981).
2. Graff, Y. C., Wohlhueter, R. M. & Plagemann, P. G. W. *J. cell. Physiol.* **96**, 171–188 (1978).
3. Wohlhueter, R. M. & Plagemann, P. G. W. *Int. Rev. Cytol.* **64**, 171–240 (1980).
4. Imprim, C. C., Foster, K., Micklem, K. J. & Pasternak, C. A. *Biochem. J.* **186**, 847–860 (1980).
5. Hassell, J. A., Colby, C. & Romano, A. H. *J. cell. Physiol.* **86**, 37–45 (1975).

WHITE *ET AL.* REPLY—Since our paper¹ was submitted, we have succeeded in establishing conditions in which the

uptake of 3-O-methyl-D-glucose in our cells is sufficiently linear to permit formal kinetic analysis. This has enabled us to repeat our experiments using a glucose analogue that is not phosphorylated. The V_{max} values for 3-O-methyl-D-glucose uptake are comparable with those for 2-deoxyglucose uptake, but the K_m values are two to four times higher, indicating that 3-O-methyl-D-glucose is transported less efficiently than 2-deoxyglucose or glucose. Nonetheless, the reduction in K_m that we previously described for 2-deoxyglucose uptake also occurs with 3-O-methyl-D-glucose.

Table 1 gives representative values for CBAT₆T₆ mouse fibroblasts, PG19 malignant mouse melanoma cells, 1Acn2, a human hybrid cell line in which malignancy is suppressed and 1Acn1TG, a malignant segregant derived from this hybrid. The reduction in K_m that we described previously¹ is thus not primarily determined by the phosphorylation reaction, but by the flux of hexose across the cell membrane. Measurements we have made of the hexokinase activity in cell homogenates indicate that there is no correlation between hexokinase activity and malignancy.

Our conclusion that malignancy is closely linked to a decrease in the K_m of the membrane hexose transport system stands.

M. K. WHITE
M. E. BRAMWELL
H. HARRIS

Sir William Dunn School
of Pathology,
University of Oxford,
South Parks Road,
Oxford OX1 3RE, UK

1. White, M. K., Bramwell, M. E. & Harris, H. *Nature* **294**, 232–235 (1981).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 500 words and the briefest of replies, to the effect that a point is taken, should be considered.

BOOKS RECEIVED

Biological Sciences

- GIBBS, E.P.J. (ed.). *Virus Diseases of Food Animals. A World Geography of Epidemiology and Control*. Vol. II, Disease Monographs. Pp.456. ISBN 0-12-282202-1. (Academic: 1981.) £33.80, \$69.50.
- GILBERT, D.L. (ed.). *Oxygen and Living Processes: An Interdisciplinary Approach*. Pp.401. ISBN 3-540-90554-5. (Springer-Verlag: 1981.) Np.
- GODFRAIND, T. and MEYER, P. (eds). *Cell Membrane in Function and Dysfunction of Vascular Tissue. Proceedings of the 5th Argenteuil Symposium held in Waterloo, Belgium, June-July 1980*. Pp.270. ISBN 0-444-80316-5. (Elsevier/North-Holland Biomedical: 1981.) \$74.50, Dfl.175.
- GULL, K. and OLIVER, S.G. (eds). *The Fungal Nucleus. Symposium of The British Mycological Society held at Queen Elizabeth College, London, April 1980*. Pp.358. ISBN 0-521-23492-1. (Cambridge University Press: 1981.) £32.50.
- HAMILTON, A. *A Field Guide to Uganda Forest Trees*. Pp.279. (no ISBN) (The New University of Ulster. Obtainable from Dr A.C. Hamilton, Botany Department, Makerere University, P.O. Box 7062, Kampala, Uganda: 1981.) Pbk (Uganda) 400/-shs. (elsewhere) £12.50.
- HERREID, C.F. II and FOURTNER, C.R. (eds). *Locomotion and Energetics in Arthropods. Proceedings of a Symposium held December 1980 at the University of Washington, Seattle, Washington*. Pp.546. ISBN 0-306-40830-9. (Plenum: 1981.) \$59.50.
- JAMIESON, B.G.M. *The Ultrastructure of the Oligochaeta*. Pp.462. ISBN 0-12-380180-X. (Academic: 1981.) £68, \$139.50.
- KANNO, T. (ed.). *Paraneurons: Their Features and Function. Proceedings of the International Symposium, Shikotsuko, Hokkaido, Japan, July 1979*. Pp.193. ISBN 90-219-0479-9. (Excerpta Medica, Amsterdam: 1981.) \$38.25, Dfl.90.
- KITCHING, R.L. and JONES, R.E. *The Ecology of Pests: Some Australian Case Histories*. Pp.254. Pbk ISBN 0-643-00408-4. (CSIRO, East Melbourne, Victoria, Australia: 1981.) Pbk \$A 8.50.
- KOORYMAN, G.L. *Weddell Seal Consummate Diver*. Pp.135. ISBN 0-521-23657-6. (Cambridge University Press: 1981.) £18.
- KURSTAK, E. and KURSTAK, C. (eds). *Comparative Diagnosis of Viral Diseases. Vol. IV, Vertebrate Animal and Related Viruses Part B — RNA Viruses*. Pp.694. ISBN 0-12-429704-8. (Academic: 1981.) \$79.
- LSAT, F.T. and GARDINER, A.S. (eds). *NERC Institute of Terrestrial Ecology. Forest and Woodland Ecology. An Account of Research Being Done in ITE. ITE Symposium, No.8*. Pp.158. Pbk ISBN 0-904282-51-1. (Institute of Terrestrial Ecology, Cambridge: 1981.) Pbk £5.
- LUMSDEN, W.H.R., MULLER, R. and BAKER, J.R. (eds). *Advances in Parasitology, Vol.19*. Pp.226. ISBN 0-12-031719-2. (Academic: 1981.) £16.40, \$39.50.
- MABBERLEY, D.J. (ed.). *Revolutionary Botany 'Thalassiphyta' and Other Essays of A.H. Church. (with a 'Recollection' by E.J.H. Corner)*. Pp.256. ISBN 0-19-854548-7. (Oxford University Press: 1981.) £21.
- MAJKOWSKI, J. (ed.). *Epilepsy. A Clinical and Experimental Research. Monographs in Neural Sciences, Vol.5. Proceedings of the 2nd European Regional Conference, Warsaw, October 1978*. Pp.304. Flexi ISBN 3-8055-0635-X. (Karger, Basel: 1980.) SwFr. 94, DM 113, \$56.50.
- MATTHEWS, C.D. and SEAMARK, R.F. (eds). *Pineal Function. Proceedings of the Satellite Symposium 6th International Congress of Endocrinology, Melbourne, Australia, February 1980*. Pp.271. ISBN 0-444-80313-0. (Elsevier/North-Holland Biomedical: 1981.) £68.90, Dfl.160.
- MCANARNEY, E.R. and STICKLE, G. (eds). *Pregnancy and Childbearing During Adolescence: Research Priorities for the 1980s. March of Dimes Birth Defects Foundation: Original Article Series, Vol.XVII, Number 3, 1981*. Pp.176. ISBN 0-8451-1043-8. (Alan R. Liss, New York: 1981.) £19.10, DM 84.
- MCDONALD, J.N. *North American Bison: Their Classification and Evolution*. Pp.316. ISBN 0-520-04002-3. (University of California Press: 1981.) £24.50, \$43.70.
- McLAREN, A. *Germ Cells and Soma: A New Look at an Old Problem*. Pp.119. ISBN 0-300-02694-3. (Yale University Press: 1981.) £10.50.
- MINDERHOUD, J.M. (ed.). *Cerebral Blood Flow. Basic Knowledge and Clinical Implications. The Jonxis Lectures. A series of post-graduate medical courses, Vol.7*. Pp.313. ISBN 90-219-6007-9. (Excerpta Medica, Amsterdam: 1981.) \$36.50, Dfl.86.
- MINORS, D.S. and WATERHOUSE, J.M. *Circadian Rhythms and the Human*. Pp.332. ISBN 0-7236-0592-0. (John Wright, Bristol: 1981.) £15.
- NIEUWKOOP, P.D. and SUTASURYA, L.A. *Primordial Germ Cells in the Invertebrates From Epigenesis to Preformation*. Pp.258. ISBN 0-521-22189-7. (Cambridge University Press: 1981.) £30.
- PROKOP, R. *Fossils. A Hamlyn Colour Guide*. Pp.224. ISBN 0-600-35370-2. (Hamlyn, Feltham, UK: 1981.) £2.95.
- RODRIGO, R.G.A. *The Alkaloids. Chemistry and Physiology, Vol. XVIII*. Pp.411. ISBN 0-12-469518-3. (Academic: 1981.) \$65.
- SARMA, R.H. (ed.). *Biomolecular Stereodynamics, Vol. 1. Proceedings of a Symposium held at the State University of New York at Albany, April 1981*. Pp.472. ISBN 0-940030-00-4. (Adenine Press, New York: 1981.) Np.
- SARMA, R.H. (ed.). *Biomolecular Stereodynamics, Vol.2. Proceedings of a Symposium held at the State University of New York at Albany, April 1981*. Pp.501. ISBN 0-940030-01-2. (Adenine Press, New York: 1981.) Np.
- SATO, V.L. and GEFTER, M.L. (eds). *Cellular Immunology. Selected Readings and Critical Commentary*. Pp.647. Pbk ISBN 0-201-10434-2. (Addison-Wesley: 1981.) Pbk \$21.50.
- SCHOCKAERT, E.R. and BALL, I.R. (eds). *The Biology of the Turbellaria. Developments in Hydrobiology, 6. Proceedings of the 3rd International Symposium held in Diepenbeek, Belgium*. Pp.300. ISBN 90-6193-757-4. (W. Junk, The Hague: 1981.) Np.
- SCHULMAN, J.D. and SIMPSON, J.L. (eds). *Genetic Diseases in Pregnancy. Maternal Effects and Fetal Outcome*. Pp.493. ISBN 0-12-630940-X. (Academic: 1981.) Np.
- SILBER, R. *et al.* (eds). *Contemporary Hematology/Oncology, Vol. 2*. Pp.548. ISBN 0-306-40683-7. (Plenum Medical: 1981.) \$49.50.
- SOKAL, R.R. and ROHLF, F.J. *Biometry: The Principles and Practice of Statistics in Biological Research, 2nd Edn*. Pp.859. ISBN 0-7167-1254-7. (W.H. Freeman: 1981.) £19.50.
- SPEEDING, C.R.W., WALSHINGHAM, J.M. and HOXEY, A.M. *Biological Efficiency in Agriculture*. Pp.369. ISBN 0-12-656560-0. (Academic: 1981.) £18.40, \$44.50.
- SPENCER, D.M. (ed.). *The Downy Mildews*. Pp.636. ISBN 0-12-656860-X. (Academic: 1981.) £49.60, \$119.50.
- SQUIRE, J. *The Structural Basis of Muscular Contraction*. Pp.698. ISBN 0-306-40582-2. (Plenum: 1981.) \$65.
- STEWART, A.D. and HUNT, D.M. *The Genetic Basis of Development. Tertiary Level Biology*. Pp.221. Hbk ISBN 0-216-91161-3; pbk ISBN 0-216-91160-5. (Blackie: 1981.) Hbk £18.95; pbk £8.95.
- SYKOVA, E., HNIK, P. and VYKLICKÝ, L. (eds). *Ion-Selective Microelectrodes and Their Use in Excitable Tissues. Proceedings of a Satellite Symposium to the 28th International Congress of Physiological Sciences, held July 1980, in Prague, Czechoslovakia*. Pp.369. ISBN 0-306-40723-X. (Plenum: 1981.) \$39.50.
- THEWS, G. and VAUPEL, P. *Grundriss der Vegetativen Physiologie. Heidelberger Taschenbücher. Basistext Medizin, Biologie, Band 210. Seiten 452*. ISBN 3-540-10631-6. (Springer-Verlag: 1981.) DM 29.80, \$13.60.
- THOMPSON, T., DEWS, P.B. and MCKIM, W.A. (eds). *Advances in Behavioral Pharmacology, Vol. 3*. Pp.217. ISBN 0-12-004703-9. (Academic: 1981.) \$31.
- WALK, R.D. *Perceptual Development*. Pp.224. Pbk ISBN 0-8185-0447-1. (Brooks/Cole, Monterey, California: 1981.) Pbk \$8.95.
- WARREN, K.S. (ed.). *Coping with the Biomedical Literature: A Primer for the Scientist and the Clinician*. Pp.233. Hbk ISBN 0-03-057036-0; pbk ISBN 0-03-057034-4. (Praeger: 1981.) Hbk £22; pbk np.
- WARWICK, W.J. (ed.). *1,000 Years of Cystic Fibrosis. Collected Papers Presented at an International Conference at the University of Minnesota, St. Paul Campus, May 1981*. Pp.307. (no ISBN) (University of Minnesota: 1981.) Pbk np.
- WATSON, L. *Sea Guide to Whales of the World*. Pp.302. ISBN 0-525-93202-X. (Elsevier-Dutton, New York: 1981.) \$31 until 1/1/82; \$36.25 thereafter.
- WEST, J.B. (ed.). *High Altitude Physiology. Benchmark Papers in Human Physiology, Vol. 15*. Pp.461. ISBN 0-87933-388-X. (Academic: 1981.) \$55.
- WILKINSON, A.W. *Investigation of Brain Function. Ettore Majorana International Sciences. (Life Sciences, Vol. 7.) Proceedings of the 11th Course of the International School of Medical Sciences held May 1980, at the Ettore Majorana Centre for Scientific Culture, Erice, Sicily, Italy*. Pp.262. ISBN 0-306-40811-2. (Plenum: 1981.) \$35.

Applied Biological Sciences

- CORNELIUS, C.E. and SIMPSON, C.F. (eds). *Advances in Veterinary Science and Comparative Medicine, Vol.25*. Pp.382. ISBN 0-12-039225-9. (Academic: 1981.) \$49.
- DOMINO, E.F. and LAFAYETTE CLINIC. *PCP (Phencyclidine): Historical and Current Perspectives*. Pp.537. ISBN 0-916182-03-7. (NPP Books, Ann Arbor, Michigan: 1981.) \$40.
- GAHL, G.M., KESSEL, M. and NOLPH, K.D. (eds). *Advances in Peritoneal Dialysis. Proceedings of the 2nd International Symposium, Berlin (-West), June 1981*. Pp.508. ISBN 90-219-9537-9. (Excerpta Medica, Amsterdam: 1981.) \$83.75, Dfl. 180.
- LADER, M.H. and RICHENS, A. (eds). *Methods in Clinical Pharmacology — Central Nervous System*. Pp.164. ISBN 0-333-31726-2. (Macmillan Press, London & New York: 1981.) £20.
- MANCIAUX, M. (ed.). *Handicaps in Childhood. Child Health and Development, Vol. 1*. Pp.182. ISBN 3-8055-2935-X. (Karger, Basel: 1982.) SwFr. 83, DM 99, \$49.75.
- MCDONALD, J.C. (ed.). *Recent Advances in Occupational Health, 1*. Pp.292. Pbk ISBN 0-443-02300-X. (Churchill Livingstone: 1981.) Pbk £16.
- MORSELLI, P.L. *et al.* (eds). *Neurotransmitters, Seizures, and Epilepsy*. Pp.361. ISBN 0-89004-753-7. (Raven: 1981.) \$42.
- OLD, R.W. and PRIMROSE, S.B. *Principles of Gene Manipulation: An Introduction to Genetic Engineering, 2nd Edn. Studies in Microbiology, Vol. 2*. Pp.214. Pbk ISBN 0-632-00856-3. (Blackwell Scientific: 1981.) Pbk £6.50.
- RACHIE, K.O. and LYMAN, J.M. (eds). *Genetic Engineering for Crop Improvement. A Rockefeller Foundation Conference, May 1980*. Pp.254. (No ISBN) (The Rockefeller Foundation: 1981.) Pbk np.
- YETIV, J.Z. and BIANCHINE, J.R. (eds). *Recent Advances in Clinical Therapeutics. Vol. 1, Hypertension, Cardiovascular Disease, Analgesics, and Endocrine Disorders*. Pp.265. ISBN 0-12-788950-7. (Academic: 1981.) \$29.50.

Psychology

- ARKIN, A.M. *Sleep-Talking: Psychology and Psychophysiology*. Pp.624. ISBN 0-89859-031-0. (Lawrence Erlbaum, London: 1981.) £19.95.
- BAN, T.A. *Psychopharmacology of Depression: A Guide for Drug Treatment*. Pp.130. ISBN 3-8055-1154-X. (Karger, Basel: 1980.) SwFr. 32, DM 38, \$19.25.

EMRICH, H.M. (ed.). *The Role of Endorphins in Neuropsychiatry. Modern Problems of Pharmacopsychiatry*, Vol. 16. Pp.292. ISBN 3-8055-2918-X. (Karger, Basel: 1981.) SwFr. 133, DM 159, \$79.75.

GOLEMAN, D. and FREEDMAN, J. *What Psychology Knows that Everyone Should*. Pp.270. ISBN 0-86616-010-8. (Stephen Greene, Brattleboro, Vermont: 1981.) \$12.95.

GOODWIN, C. *Conversational Organization. Interaction between Speakers and Hearers. Language, Thought and Culture: Advances in the Study of Cognition*. Pp.195. ISBN 0-12-289780-3. (Academic: 1981.) \$26.

JEFFERSON, J.W. and MARSHALL, J.R. *Neuropsychiatric Features of Medical Disorders*. Pp.383. ISBN 0-306-40674-8. (Plenum Medical: 1981.) \$29.50.

MATSON, J.L. and MCCARTNEY, J.R. (eds). *Handbook of Behavior Modification with the Mentally Retarded*. Pp.414. ISBN 0-306-40617-9. (Plenum: 1981.) \$39.50.

OLLENDICK, T.H. and CERNY, J.A. *Clinical Behavior Therapy with Children*. Pp.351. ISBN 0-306-40774-4. (Plenum: 1981.) \$24.95.

SARTORIUS, N. and HELMCHEN, H. (eds). *Multicentre Trials. Modern Problems of Pharmacopsychiatry*, Vol. 16. Pp.116. ISBN 3-8055-2806-X. (Karger, Basel: 1981.) SwFr. 79, DM 95, \$47.50.

WHEATLEY, D. (ed.). *Psychopharmacology of Sleep*. Pp.256. ISBN 0-89004-593-3. (Raven: 1981.) \$29.

WILKE, J.T. *A Neuropsychological Model of Knowing*. Pp.79. Hbk ISBN 0-8191-1768-4; pbk ISBN 0-8191-1769-2. (University Press of America, Washington: 1981.) Hbk \$16.50; pbk \$6.75.

Sociology

FELDBRUGGE, J.T.T.M. and VONELGG, Y.A.W. (eds). *Involuntary Institutionalization. Changing Concepts in the Treatment of Delinquency. Proceedings of a Workshop held on the occasion of the 25th Anniversary of the Dr. Henri van der Hoeven Kliniek, Utrecht, The Netherlands, April 1980*. Pp.106. ISBN 90-219-0507-8. (Excerpta Medica, Amsterdam: 1981.) \$21.25, Dfl. 50.

SCHACTER, A.J. (ed.). *Drug Dependence and Alcoholism. Vol. 2, Social and Behavioral Issues. 2nd half of the Proceedings of the 1978 National Drug Abuse Conference, held in Seattle, Washington, April 1978*. Pp.1041. ISBN 0-306-40324-2. (Plenum: 1981.) \$39.75.

SIEBER, S.D. *Fatal Remedies. The Ironies of Social Intervention*. Pp.234. ISBN 0-306-40717-5. (Plenum: 1981.) \$24.50.

Anthropology

AUGÉ, M. *The Anthropological Circle. Symbol, Function, History*. Pp.131. Hbk ISBN 0-521-23236-8; pbk ISBN 0-521-28548-8. (Cambridge University Press: 1982.) Hbk £12.50; pbk £4.95.

LANGHAM, I. *The Building of British Social Anthropology. W.H.R. Rivers and his Cambridge Disciples in The Development of Kinship Studies, 1898-1931. Studies in the History of Modern Science*, 8. Pp.380. Hbk ISBN 90-277-1264-6; pbk ISBN 90-277-1265-4. (Riedel: 1981.) Hbk Dfl. 145, \$72.50; pbk Dfl. 65, \$29.50.

MUIZON, C. de. *Recherche sur les Grandes Civilisations. Mémoire No. 6, Les Vertébrés Fossiles de la Formation Pisco (Pérou). Institut Français D'Études Andines*. Pp.150. Pbk ISBN 2-86538-013-0. (Editions A.D.P.F., Paris: 1981.) Pbk np.

SCHUYLER, R.L. *Archaeological Perspectives on Ethnicity in America. Afro-American and Asian American Culture History. Baywood Monographs in Archaeology*, 1. Pp.147. Pbk ISBN 0-89503-018-7. (Baywood Farmingdale, New York: 1980.) Pbk np.

SHACKLEY, M. *Environmental Archaeology*. Pp.256. Hbk ISBN 0-04-913020-X; pbk ISBN 0-04-913021-8. (Allen & Unwin, London: 1981.) Hbk £18; pbk £9.95.

Education

BROWN, L. (ed.). *Sex Education in the Eighties. The Challenge of Healthy Sexual Evolution. Perspectives in Sexuality Behavior, Research and Therapy*. Pp.264. ISBN 0-306-40762-0. (Plenum: 1981.) \$24.50.

INGLE, R. and JENNINGS, A. *Science in Schools: Which Way Now? Studies in Education. A Series of Occasional Papers*, No.8. Pp.185. Pbk ISBN 0-85473-100-8. (University of London Institute of Education: 1981.) Pbk £5.30.

RHINE, W.R. (ed.). *Making Schools More Effective*. Pp.360. ISBN 0-12-587060-4. (Academic: 1981.) \$28.50.

ROSENBLUM, S. and LOUIS, K.S. *Stability and Change: Innovation in an Educational Context*. Pp.352. ISBN 0-306-40665-9. (Plenum: 1981.) \$29.50.

General

ADELMAN, W.J. Jr. and GOLDMAN, D.E. (eds). *The Biophysical Approach to Excitable Systems. A Volume in Honour of Kenneth S. Cole on his 80th Birthday*. Pp.258. ISBN 0-306-40784-1. (Plenum: 1981.) \$29.50.

AUDOUZE, J. *Aujourd'hui L'Univers*. Pp.343. Pbk ISBN 2-7-144-1423-0. (Pierre Belfond, Paris: 1981.) Pbk np.

BARNEA, J., MEYER, R.F. and OLSON, J.C. (eds). *Long-Term Energy Resources, Vol. I. An International Conference sponsored by UNITAR, November-December 1979, Montreal, Quebec, Canada*. Pp.729. ISBN 0-273-08534-4. (Pitman Books: 1981.) £30 per Vol; £80 3-Vol. set.

BARNEA, J., MEYER, R.F. and OLSON, J.C. (eds). *Long-Term Energy Resources, Vol. II. An International Conference sponsored by UNITAR, November-December 1979, Montreal, Quebec, Canada*. Pp.745. ISBN 0-273-08534-4. (Pitman Books: 1981.) £30 per Vol; £80 3-Vol. set.

BARNEA, J., MEYER, R.F. and OLSON, J.C. (eds). *Long-Term Energy Resources, Vol. III. An International Conference sponsored by UNITAR, November-December 1979, Montreal, Quebec, Canada*. Pp.630. ISBN 0-273-08534-4. (Pitman Books: 1981.) £30 per Vol; £80 3-Vol. set.

BENNETT, R.J. (ed.). *European Progress in Spatial Analysis*. Pp.305. ISBN 0-85086-1. (Methuen: 1981.) £12.50.

BLACK, P.E. *Environmental Impact Analysis*. Pp.146. ISBN 0-03-059618-1. (Praeger, a Division of CBS Inc, New York: 1981.) Np.

BRINCKMAN, F.E. and FISH, R.H. (eds). *Environmental Speciation and Monitoring Needs for Trace Metal Containing Substances from Energy-Related Processes. Proceedings of the DoE/NBS Workshop held at the National Bureau of Standards, Gaithersburg, Maryland, May 1981*. Pp.324. (no ISBN) (US Department of Commerce, National Bureau of Standards, Washington: 1981.) Pbk np.

BURBIDGE, G. and HEWITT, A. (eds). *Telescopes for the 1980s*. Pp.278. ISBN 0-8243-2902-3. (Annual Reviews, Palo Alto, California: 1981.) Np.

CHICHESTER, C.O., MARK, E.M. and STEWART, G.F. (eds). *Advances in Food Research*, Vol. 27. Pp.366. ISBN 0-12-016427-2. (Academic: 1981.) \$39.

DE NEUFVILLE, J.I. *The Land Use Policy Debate in the United States*. Pp.269. ISBN 0-306-40718-3. (Plenum: 1981.) \$29.50.

EDITORS OF SMITHSONIAN. *The Best of Smithsonian*. Pp.320. ISBN 0-517-54526-8. (Harmony Books, New York: 1981.) \$27.50.

FRASER, J.T., LAWRENCE, N. and PARK, D. (eds). *The Study of Time, IV. Papers from the 4th Conference of the International Society for the Study of Time, Alpbach, Austria*. Pp.286. ISBN 3-540-90594-4. (Springer-Verlag: 1981.) DM 98, \$45.70.

FRY, B.M. and HERNON, P. *Government Publications: Key Papers. Guides to Official Publications*, Vol. 8. Pp.814. ISBN 0-08-025216-8. (Pergamon: 1981.) £42, \$100.

GERLACH, S.A. *Marine Pollution. Diagnosis and Therapy*. Pp.218. ISBN 3-540-10940-4. (Springer-Verlag: 1981.) DM 58, \$24.70.

GRAHAM, L.R. *Between Science and Values*. Pp.449. ISBN 0-231-05192-1. (Columbia University Press: 1981.) \$25.90.

GREIDANUS, J.H. *A Solution of the Problem of Mind and Matter*. Pp.45. Pbk ISBN 04448-5527-0. (North-Holland: 1981.) Pbk np.

HALL, L. *Earth's Song*. Pp.274. ISBN 0-8262-0323-X. (University of Missouri Press: 1981.) £8.95.

HAND, D.J. *Discrimination and Classification*. Pp.218. ISBN 0-471-28048-8. (Wiley: 1981.) £15.50, \$36.75.

JONES, M. Jr. and MOSS, R.A. (eds). *Reactive Intermediates. A Serial Publication*, Vol. 2. Pp.396. ISBN 0-471-01875-9. (Wiley: 1981.) £38.85, \$69.85.

KAUFMAN, K. *Of Trees, Leaves, and Ponds. Studies in Photo-Impressionism*. Pp.64. Pbk ISBN 0-525-47678-4. (Dutton, New York: 1981.) Pbk \$11.50.

LAURENCE, J.C. and NOBLE, A.G. (eds). *The Environment: Chinese and American Views*. Pp.397. ISBN 0-416-32050-3. (Methuen: 1981.) £16.50.

MARKLEY, O.W. and HARMAN, W.W. (eds). *Changing Images of Man*. Pp.255. Hbk ISBN 0-08-024314-2; pbk ISBN 0-08-024313-4. (Pergamon: 1981.) £7.95, \$17.50.

McMULLAN, J.T. and MORGAN, R. *Heat Pumps*. Pp.156. ISBN 0-85274-419-6. (Adam Hilger, Bristol: 1981.) £15.

MCRAE, S.G. and BURNHAM, C.P. *Land Evaluation*. Pp.239. ISBN 0-19-854518-5. (Clarendon Press, Oxford: 1981.) £18.

NEILSON, B.J. and CRONIN, L.E. (eds). *Estuaries and Nutrients*. Pp.641. ISBN 0-89603-035-0. (Human Press, Clifton, New Jersey: 1981.) \$64.50 (US); \$74.50 (elsewhere).

NEMETZ, P.N. (ed.). *Energy Crisis. Policy Response*. Pp.304. Pbk ISBN 0-920380-49-2. (Institute for Research on Public Policy, Montreal: 1981.) Pbk np.

O'RIORDAN, T. and TURNER, R.K. (eds). *Progress in Resource Management and Environmental Planning*, Vol.3. Pp.324. ISBN 0-471-27968-4. (Wiley: 1981.) £23.75, \$59.

PERROT, J. and VANDERMEERSCH, B. (eds). *Paléorient*, Vol. 6. Pp.319. Pbk ISBN 2-222-02828-0. (CNRS, Paris: 1981.) Pbk np.

PRINGLE, P. and SPIGELMAN, J. *The Nuclear Barons*. Pp.578. ISBN 0-7181-2061-2. (Michael Joseph, London: 1982.) £12.95.

RODWIN, L. *Cities and City Planning. Environment, Development and Public Policy Series*. Pp.309. ISBN 0-306-40666-7. (Plenum: 1981.) \$29.50.

RONAN, C.A. *The Shorter Science and Civilisation in China. An Abridgement of Joseph Needham's Original Text*. Vol. 2, Vol.3 and a section of Vol.4, Part 1 of the major series. Pp.459. ISBN 0-521-23582-0. (Cambridge University Press: 1981.) £15.

ROTLAT, J. *Nuclear Radiation in Warfare*. Pp.149. ISBN 0-85066-217-6. (Taylor & Francis, London: 1981.) £9.50, \$20.

SABOM, M.B. *Recollections of Death. A Medical Investigation*. Pp.224. ISBN 0-06-014895-0. (Harper & Row: 1982.) \$13.50.

SALVENDY, G. and SMITH, M.J. (eds). *Machine Pacing and Occupational Stress. Proceedings of the International Conference, Purdue University, March 1981*. Pp.374. ISBN 0-85066-225-7. (Taylor & Francis, London: 1981.) £19.75.

TRUPP, F. *The Last Indians: South America's Cultural Heritage*. Pp.257. ISBN 3-85399-023-1. (Prentice-Hall: 1981.) £14.95.

ULRICH, G.E., HODGES, C.A. and MUEHLBERGER, W.R. (eds). *Geology of the Apollo 16 Area: Central Lunar Highlands. Geological Survey Professional Paper*, 1048. Pp.539+plates. (no ISBN) (US Government Printing Office, Washington: 1981.) Pbk np.

UNWIN, D. *Introductory Spatial Analysis*. Pp.212. Hbk ISBN 0-416-72190-7; pbk ISBN 0-416-72200-8. (Methuen: 1981.) Hbk £11.50; pbk £5.50.

WEST, E. *The Kitchen in the Hills. The Hovel Cookbook*. Pp.184. ISBN 0-571-11709-0. (Faber & Faber: 1981.) £6.25.

WHITFIELD, M. and JAGNER, D. (eds). *Marine Electrochemistry: A Practical Introduction*. Pp.529. ISBN 0-471-27976-5. (Wiley: 1981.) £27.50, \$65.

WOLF, J.B. *Fear of Fear. A Survey of Terrorist Operations and Controls in Open Societies. Criminal Justice and Public Safety*. Pp.235. ISBN 0-306-40766-3. (Plenum: 1981.) \$22.